Renewable and Sustainable Energy Reviews 26 (2013) 108–134

Contents lists available at ScienceDirect

Renewable and Sustainable Energy Reviews

journal homepage: www.elsevier.com/locate/rser

A review of LCA greenhouse gas emissions results for advanced biofuels: The use of meta-regression analysis

Fabio Menten [a,b,*], Benoît Chèze [a,c], Laure Patouillard [a], Frédérique Bouvart [a]

[a] IFP Energies nouvelles, nouvelles, 1 et 4 Avenue de Bois-Préau, 92852 Rueil-Malmaison, France
[b] Arts et Métiers ParisTech, Esplanade des Arts et Métiers, 33405 Talence, France
[c] EconomiX-CNRS, University of Paris Ouest, UMR210 Economie Publique, INRA, Climate Economics Chair, Paris, France

ARTICLE INFO

Article history:
Received 6 December 2012
Received in revised form
17 April 2013
Accepted 20 April 2013
Available online 19 June 2013

Keywords:
Biofuels
GHG
LCA
Meta-analysis

ABSTRACT

This article presents the results of a literature review performed with a meta-regression analysis (MRA). It focuses on the estimates of advanced biofuel Greenhouse Gas (GHG) emissions determined with a Life Cycle Assessment (LCA) approach. The mean GHG emissions of both second (G2) and third generation (G3) biofuels and the effects of factors influencing these estimates are identified and quantified by means of specific statistical methods. 47 LCA studies are included in the database, providing 593 estimates. Each study estimate of the database is characterized by (i) technical data/characteristics, (ii) author's methodological choices and (iii) typology of the study under consideration. The database is composed of both the vector of these estimates—expressed in grams of CO_2 equivalent per MJ of biofuel (g CO_2eq/MJ) and a matrix containing vectors of predictor variables which can be continuous or dummy variables. The former is the dependent variable while the latter corresponds to the explanatory variables of the meta-regression model. Parameters are estimated by means of econometrics methods.

Our results clearly highlight a hierarchy between G3 and G2 biofuels: life cycle GHG emissions of G3 biofuels are statistically higher than those of Ethanol which, in turn, are higher than those of BtL. Moreover, this article finds empirical support for many of the hypotheses formulated in narrative literature surveys concerning potential factors, which may explain estimates variations. Finally, the MRA results are used to address the harmonization issue in the field of advanced biofuels GHG emissions thanks to the technique of *benefits transfer using meta-regression models*. The range of values hence obtained appears to be lower than the fossil fuel reference (about 83.8 in g CO_2eq/MJ). However, only Ethanol and BtL do comply with the GHG emission reduction thresholds for biofuels defined in both the American and European directives.

Contents

* Corresponding author at: IFP Energies Nouvelles, 1 et 4 Avenue de Bois-Préau, 92852 Rueil-Malmaison, France. Tel.: +33 9 5174 6017; fax: +33 1 4752 7082.
E-mail address: fabio.menten@gmail.com (F. Menten).

1364-0321/$ - see front matter
http://dx.doi.org/10.1016/j.rser.2013.04.021

1. Introduction

This article addresses the environmental evaluation issues of advanced biofuels in the transport sector. It focuses on a specific environmental evaluation method—Life Cycle Assessment (LCA) and its estimates of second (G2) and third generation (G3) biofuels greenhouse gas (GHG) emissions. The mean Global Warming impact indicator, expressed in grams of CO_2 equivalent per MJ of biofuel (g CO_2eq/MJ), and the effects of factors influencing these estimates are characterized and quantified using a *meta-regression analysis* (MRA): a quantitative research method to review and synthesize empirical literature. This research is of primary importance as this measure may be interpreted as an estimate of the contribution to climate change of advanced biofuels.

The transport sector is unique because it relies almost exclusively on oil, which represented 94% of all transportation fuels in 2011 [1]. In the current context of rising oil prices associated with concerns about global warming and energy security, alternative transportation fuels, such as biofuels, are being developed. They are viewed as a feasible and sustainable alternative to petroleum based fuels. This paper focuses on liquid biofuels, which can be used without major modifications in current engines in the transport sector.

First generation liquid biofuels (thereafter named as G1 biofuels) are economically viable and produced in industrial scale nowadays mainly from crops such as sugar cane, sugar beet, wheat, corn, rapeseed, sunflower, etc. Ethanol and biodiesel are the most representative categories of these biofuels. These G1 biofuels have come up against sustainability issues mostly related to the use of agricultural commodities in their production processes. Indeed, the production of G1 biofuels induces an additional demand for cultivated plants and, consequently, an increased use of arable land. Furthermore, it has been suggested that it may induce a rise of food prices [2]. Additionally, many life-cycle based studies point out that G1 biofuels do not reduce GHG emissions as significantly as expected or have a low net energy output [3]. As a consequence, G2 and G3 liquid biofuels from biomass residues, non-alimentary crops and wastes have been developed in the recent years. These biofuels seem to be more efficient than G1 biofuels in terms of land use, food security, GHG emission reductions and other environmental aspects [4].

G2 Ethanol is obtained from the biochemical conversion of lignocellulosic biomass[1] . Synthetic diesel from biomass, also known as Biomass to Liquids (BtL) or biomass FT-diesel, is produced by the thermochemical conversion of lignocellulosic biomass. In this paper, G2 biofuels refers to both of these biofuels. G3 biofuels are produced from microalgae using algal oil for biodiesel production from conventional transesterification (a.k.a Fatty Acid Methyl Ester, FAME) or hydrotreated algal oil (HAO). The cited G2 and G3 biofuels are referred to in this paper as advanced biofuels[2] (see Appendix A for further details on their production processes). They are currently either in research and development or demonstration phase and still need further improvements to be commercially viable.

Some states have set ambitious production targets for biofuels, supported by subsidies and legislative incentives. In the European Union (EU), the Renewable Energy Directive (RED, [5]) requires the use of 10% of renewable energies in the transport sector by 2020 (in 2009, the share was 3.6%). To achieve this goal, the contribution of biofuels produced from lignocellulosic materials, wastes and residues is considered to be twice that made by other biofuels. This can be viewed as an incentive for the development of advanced biofuels. In the United States (US), the Renewable Fuel Standard (RFS2, [6]), under the US Energy Independence and Security Act of 2007, requires the use of 136 billion liters of biofuels by 2022 (in 2009, 41.9 billion litters were mandated). It specifies that 79.3 billion litters must be of “advanced biofuels” and “cellulosic biofuels” (the definition of “advanced biofuels” in the RFS2 is different from the one adopted in this paper and will be clarified later on). In addition, other countries (Australia, China, Japan, New Zealand, Brazil and others) have already been actively developing next generation biofuels and feedstock although there is little policy support in these regions [7].

Furthermore, the EU and the US set a list of sustainability requirements for biofuel production. In both regions, the only mandatory quantitative criterion is related to life cycle GHG emissions calculated using the LCA method. The RED sets minimum life cycle GHG emission savings for all biofuels compared to a fossil fuel reference. These savings are of 35% since 2009, and will be of 50% in 2017 and 60% from 2018 onward for new biofuel plants. The RFS2 also sets minimum life cycle GHG emission savings that biofuels have to comply with in order to be eligible for appropriate subsidies. Those savings are set to 20% for first generation biofuels, 50% to be considered as “advanced

[1] Lignocellulosic biomass refers to annual crop residues (e.g. corn stover), forest residues, herbaceous energy crops (e.g. switchgrass, miscanthus) and woody biomass (e.g. poplar, eucalyptus).

[2] There are different types of advanced biofuels being currently developed (methanol, dimethyl ether, butanol, hydrogen, etc.). In this paper we address only those that were the subject of a substantial number of LCA studies.

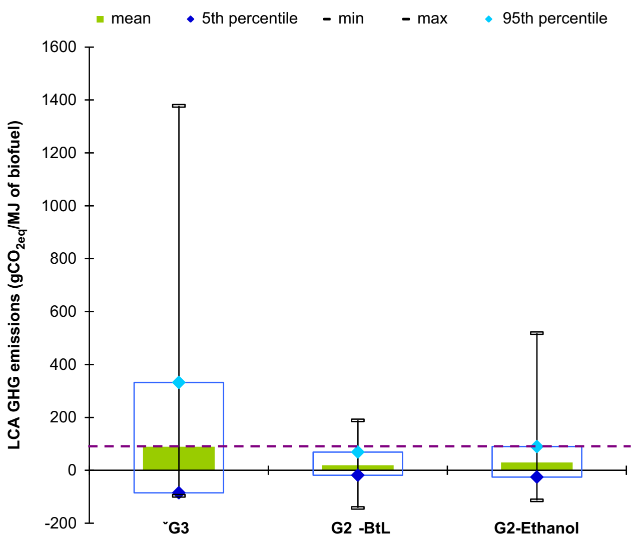


Fig. 1. GHG emissions extrema for bibliographic results of G2 and G3 biofuel LCA studies (47 studies, 593 observations).

biofuel" (as defined in the RFS2, different from our definition) and 60% to be considered as "cellulosic biofuel".

Those GHG emission requirements as well as biofuel incorporation targets are clearly in favor of G2 and G3 biofuels. This shows the will of policy makers to support their future development compared to G1 biofuels. That is one of the reasons why we choose to focus on advanced biofuels in this study.

We choose to conduct our literature analysis by reviewing only LCA studies assessing Global Warming impact indicators, i.e. GHG emissions, for the following two reasons. First, one of the main objectives for developing biofuels is to reduce global GHG emissions in order to mitigate climate change. As an illustration, recall that the only quantitative mandatory requirement for biofuel sustainability is related to life cycle GHG emission savings in the EU and in the US. Thus, it appears important to check advanced biofuel compliance with this requirement by comparing their life cycle GHG emissions with those of a fossil fuel reference. Second, a significant literature already exists that assesses GHG emissions of advanced biofuels using the LCA approach. Hence a sufficient number of studies are available to investigate this issue. Note that because GHG emissions have an environmental impact at a global scale (GHG emission effects do not depend on the place where they have been emitted), this literature review includes worldwide studies.

The first applications of LCA to biofuels to measure a Global Warming impact indicator were carried out on G1 biofuels in the 90's (such as Kaltschmitt et al. [8]). Since, numerous LCA studies were conducted to analyze G2 and G3 biofuel pathways. Despite this substantial literature, the extent to which advanced biofuels may have lower GHG emissions than the fossil reference remains a subject of debate. While the majority of these studies show GHG benefits for advanced biofuels compared to a fossil fuel reference, some authors come to the opposite conclusion. For instance, LCA GHG emission results selected for this study (47 studies providing 593 GHG emission results, see next section for more details) range from −142 (G2) to 1378 (G3) g CO_2eq/MJ of biofuel (see Fig. 1); the greatest variability of GHG emission results being for G3 biofuels.

When looking at Fig. 1, one can wonder (i) if there is a consensus about GHG emission benefits from advanced biofuels and (ii) why there is so much variation among results of these studies even though they are all investigating the same phenomenon.

Actually, even if the LCA approach is consistent throughout, each study—by nature concerns different pathways and uses specific data and methodological assumptions. Previous narrative surveys of biofuel LCA studies mention that LCA results are inconclusive regarding GHG emission performances of advanced biofuels [9–14]. According to these literature reviews, LCA GHG emission results for advanced biofuels vary significantly depending on various factors such as: the assumptions made to describe the biomass production step (model used to estimate N_2O emissions and inclusion of direct and indirect land use change), the data used to describe the biomass conversion into biofuel and the general LCA methodological choices (system boundaries, the method used to account for coproducts impacts, etc.). While these indicative results from literature reviews are really useful, primary study results remain difficult to compare because of differences in technical data or methodological choices.

As a consequence, it is quite difficult to attempt any summary and to form an accurate opinion on this topic using classical literature review methods. In particular, it seems hard to provide one GHG emission estimate appropriate for advanced biofuels.

Since most studies are inconclusive, their results may not be relevant for decision support [15]. There is a strong need for harmonization of LCA results, especially for policy makers or investors, as suggested by Heath and Mann [16] with the "*LCA harmonization project*". The purpose of harmonization, as defined by Heath and Mann, is to identify and quantify key factors that influence the environmental impacts for a technology or product in order to be more conclusive concerning its real environmental performances. At present, few studies have tried to harmonize GHG emission results from various LCA studies for advanced biofuels. For instance, Handler et al. and Liu et al. [17,18] propose to harmonize GHG emission results for G3 biofuels by normalizing their LCA models using the same methodological assumptions and generic pathways.

Although it is not possible to calculate one GHG emission estimate appropriate for all advanced biofuels, we believe it remains possible to determine central tendencies based on the distribution of previous study results. To do so, this article proposes an alternative summary to previous literature reviews, using the *meta-analysis* (MA) methodology to describe and synthesize existing estimates of the LCA GHG emissions of advanced biofuels.

MA is a quantitative research method developed to compare and/or combine outcomes of different individual quantitative studies, named primary studies, with more or less similar characteristics that can be controlled for [19]. By nature, each result from a primary study (called an estimate) may be quoted to illustrate the uncertainty of estimates. Estimates of previous studies are grouped together in a database, called meta-database, according to one or more differentiating characteristics. These estimates become then the observations, also named *effect-size* (*e-s*), of the meta-database whereas the differentiating characteristics become their potential explicative variables. In a MA framework, the *e-s* is assumed to be a function of these explicative variables; function which can be specified and assessed. When this *meta-function* is estimated by the means of multi-regression techniques, i.e. specific econometrics estimators, the MA is called a *meta-regression analysis* (MRA)[3]. This multivariate setup allowed by the meta-regression framework is very usefull in the field of literature reviews as it enables us to statistically identify and quantify—ceteris paribus the effect of the most influent characteristics on the *e-s*. Thus, compared to narrative literature reviews,

[3] So defined, MRA may be viewed as a subset of MA in the literature.

the MRA methodology—thanks to its multivariate setup gives the opportunity to test the influence of specific characteristics, after having controlled for the effect of other ones. Besides, a "*meta-regression*" framework allows to produce an estimation of the mean *e-s* weighted by the systematic influence of its main drivers. Indeed, once statistically estimated, the meta-function can be used to deduce original values of the *e-s* by specifying new values for the main drivers identified corresponding to relevant case studies. This technique of *benefits transfer using meta-regression models*, as it is named in the MA literature, may be a particularly well adapted methodology to deal with the so-called harmonization issue specific to the LCA literature.

The literature of LCA studies estimating advanced biofuels GHG emissions is now large enough to support a statistical assessment of this measure of the mean Global Warming impact indicator. The primary purpose of this MRA is to identify and quantify by statistical estimates which factors among (i) technical data/characteristics, (ii) author's methodological choices and (iii) typology of the study under consideration have an impact on variations of the GHG emission estimates. The second purpose of this MRA is to generate a distribution of the potential GHG emissions of advanced biofuels and to characterize the mean Global Warming impact indicator and its standard deviation across G2 and G3 biofuels. We investigate through an application—the potential for MRA to synthesize LCA literature by highlighting the main determinants of result variability in order to perform harmonization.

This paper is organized as follows. Section 2 is a brief summary of both LCA approach applied to biofuels and MRA methodology. Section 3 is a description of the meta-database in which the *e-s* and explanatory variables are described. Meta-regression models and the associated results are presented and analyzed in Section 4. Main conclusions and methodological discussion are presented in Section 5.

2. Methods

First, this section briefly presents the LCA approach and then summarizes how it has been used in the literature to estimate Global warming impact indicators of advanced biofuels. Second, the meta-regression methodology is briefly presented. Both sections enable a better understanding of the *e-s* and explanatory variables of the MA.

2.1. General presentation of LCA method

Life Cycle Assessment (LCA) is a method based on ISO standards 14040/14044 [20,21] aimed at assessing several potential environmental impacts of a product or a service during all of its life cycle. This approach takes into account all steps of a product's life cycle: from the extraction of natural resources necessary for its production (oil, coal, gas, etc) to its end of life or destruction ("Cradle to Grave" analysis). The LCA approach enables the characterization of potential environmental performances of a production system in order to identify potential improvements and is a relevant tool for decision makers.

The methodological framework for LCA set by international ISO standards is divided into 4 steps:

(1) Goals and scope of the study: This step deals with the definition of questions that we want to answer in the study and the final users of the results. Hence all methodological assumptions, i.e. the scope of the study (system boundaries, functional unit, method to account for coproducts, environmental impact indicators, type of data, etc) are described according to the goals of the study.
(2) Life cycle inventory: Input and output flows of matter and energy as well as emissions to the environment (air, water, soil emissions and solid wastes) included in the system are listed.
(3) Life cycle impact assessment: Inventory flows are converted into potential environmental impact categories using a characterization method. Each flow can contribute to several environmental impact categories. Impact categories and associated characterization methods are chosen in accordance with the goals and scope of the study.
(4) Interpretation of results: Results are analyzed regarding the defined goal and scope of the study.

This methodological framework is also clarified in the ILCD Handbook [22] that provides further guidance to assure consistency and quality of LCA studies.

There are two main approaches adopted in LCA studies depending on the type of questions the authors want to answer: Attributional LCA (A-LCA) and Consequential LCA (C-LCA). In an A-LCA, all the flows physically linked to the product's life cycle are included in the system's boundaries [23]. C-LCA has emerged as a modeling approach that captures impacts occurring beyond direct physical relationships assessed in A-LCA [23]. It extends the system's boundaries compared to A-LCA in order to consider market information in the life cycle inventory to assess the effects of a decision on the system [24].

LCA results can also vary from one study to another because of different sources of uncertainties. These uncertainties can be of stochastic nature (i.e. uncertainties linked to values of process data or characterization factors for example) or choice uncertainties (i.e. choice of methodological assumptions, impact assessment method, system boundaries, localization of data, etc) or lack of knowledge of studied system [22]. Uncertainties should be addressed in LCA studies by applying for instance Monte Carlo methods or by conducting sensitivity analyses.

2.1.1. Specificities of LCA applied to biofuel pathways

The first applications of LCA for the environmental evaluation of biofuels were carried out in the 90's and, since then, many methodological issues concerning this product category have been emphasized. The main specific methodological assumptions on biofuel LCA studies are:

- *System boundaries*: usually, a distinction is made between "Well To Tank" (WTT) boundaries that include all steps from the production of biomass feedstock to the transport and distribution of fuel, and "Well To Wheel" (WTW) boundaries that include the WTT steps and the fuel use (end-of-life). Infrastructures may or may not be included within the system boundaries.
- *Functional unit*: it is a measure of the function of the studied system. All LCA results from the same study should be expressed in the same functional unit to enable comparison. A usual functional unit in LCA of transportation systems is a "kilometer driven by a reference vehicle on a standard driving cycle (and assuming that generally the different fuels have a similar performance in terms of acceleration, max speed, etc.)". Another classical functional unit for assessing fuels is "the consumption of one MJ of fuel in a motor" expressed in MJ.
- *Reference system*: results of the studied system have to be compared with results of a reference system (usually a fossil fuel). This reference system has to be defined in accordance with study purposes and methodological choices; in particular it must have similar boundaries, the same functional unit and similar geographical and temporal context.

- *The method to account for coproduct*: Another classical methodological issue in LCA concerns the fact that more than one product can be produced in the studied system (called coproducts). Distributing environmental burdens among products and coproducts of a process is a controversial issue in LCA. Two types of methodology are generally applied for the multiproduct cases: the substitution method and allocation method. This last method consists in sharing proportionally the environmental impacts between products and coproducts based on physical (e.g. mass, energy) or economical characteristics of the products. With the substitution method, allocation is avoided and the burdens associated to alternative ways of producing the coproduct are subtracted from the final result. The LCA ISO standards recommend the system expansion method (also called substitution method) [25,26] but the choice of the method to account for coproducts strongly depends on the purpose of the study and on the nature of the studied system.

Biofuels use biomass as raw materials. Hence, LCA applied to biofuel pathways has to deal with some classical issues linked with the biomass production:

- *Land Use Change* (*LUC*): It refers to all changes induced by land conversion or land management changes. Direct LUC is mainly treated as the above and below ground carbon release from the conversion of forests or grasslands into agricultural land. Indirect LUC refers to all changes that occur when the increased demand for agricultural products induces land conversion in other parts of the world. It is important to note that these changes not only affect GHG emissions but other environmental aspects such as biodiversity, soil fertility, etc. Indirect LUC is the main subject of debate nowadays concerning biofuel environmental assessment, especially regarding GHG emissions [27] but there is no consensus on how to account for it in LCA methodology.
- *Nitrogen cycle*: Nitrous oxide (N_2O) field emissions are known to be the subject of controversy in the biofuel LCA world since Crutzen et al. [28] published "N_2O release from agro-biofuel production negates global warming reduction by replacing fossil fuels". There is a huge uncertainty about these emissions because they depend on local factors and this gas has a high GWP (around 300 times as much as CO_2). In a G1 biofuel LCA study conducted for the French government, the uncertainty on these emissions is estimated to be 50% [29]. To estimate these emissions, some studies use the IPCC Tier 1 methodology [30] based on the amount of nitrogen fertilizer applied in the culture. However, N_2O emissions depend on other factors such as soil characteristics and climate. Other assessment methods including these factors should provide a more accurate estimation.
- *Carbon cycle*: Considering the short-term carbon cycle, many biofuel LCA studies suppose that the amount of carbon captured by the biomass during the photosynthesis is equal to the amount of carbon released in the atmosphere during the biofuel combustion. So those studies do not take into account either the carbon stored by the biomass or the carbon releases during biofuel use, this is called the carbon-neutrality hypothesis.

2.2. General presentation of meta-analysis method

> "*Meta-analysis* (…) *is defined here as an analysis of a set of published LCA results to estimate a single or multiple impacts for a single technology or a technology category, either in a statistical sense* (e.g., *following the practice in the biomedical sciences*) *or by quantitative adjustment of the underlying studies to make them more methodologically consistent* [15]."

As stated by Brandão et al. [15], meta-analysis (MA) is nothing else than a quantitative literature review as opposed to narrative one. There are various ways of collating literature results into one mean estimate in the subsets of *MA* depending on the methodology used to synthesize literature results. In the LCA field, some authors use quantitative adjustments to recalculate LCA results after harmonizing their main methodological assumptions. A different approach is to use statistical methods to gather literature results. In this case authors can use simple descriptive statistics or go beyond by the means of multi-regression techniques, i.e. specific econometrics estimators. As a reminder, the latter subset is called meta-regression analysis (MRA).

Systematic reviews of LCA studies have gained interest due to their potential to clarify the impacts of particular products or services, producing more robust and policy-relevant results [15],[4]. Most of the published so-called LCA meta-analyses rely on a quantitative adjustment of the underlying studies [15] named "harmonization" procedure adjusting other study estimates based on "more consistent methods and assumptions" [16]. These studies typically harmonize technical parameters and methodological choices such as system boundaries, allocation procedures, impact calculation method, etc. [18,31–37]. All the cited studies aim at the reduction of the variability in calculated outcomes representing a useful starting point for more precise estimates of LCA results. One of the precursors was Farrell et al. [31] who aimed at estimating reliable values for the net energy and life-cycle GHG emissions of corn Ethanol in the US. They carry out a harmonization exercise on 6 studies, adjusting their methods and data to what the authors argue to be best practices.

The MA approach applied in this study is quite different and follows the traditional MRA practice first developed in biomedical sciences or economics. To our knowledge, Bureau et al. [38] are the only authors to use this type of approach in LCA systematic reviews. They focus their MRA on the energy balance of G1 biofuels production since they consider there is too much controversy involving life-cycle GHG estimations (due to uncertainties in the quantification of N_2O emissions from agricultural production and indirect land use change). Rather than trying to determine best estimates, they aim at identifying the variables that influence the LCA results. In the same way as their study, our results show that this methodology can be consistently applied for the identification of parameters that influence a biofuel LCA result. Moreover, we have gone further by proposing a method to predict LCA results using a meta-model. This can be seen as a harmonization method alternative to the one applied currently in LCA meta-analysis ("normalization").

The Glass' pioneering articles [39–41] in educational research are usually cited in the literature as being the first ones to propose and develop this method. Over the past three decades, MA and MRA have first been extensively applied to clinical studies in psychological and educational research and then to health sciences. It is now increasingly employed in other research fields. Since the early 1990s, this method has been gradually more and more accepted in social sciences, such as marketing and economics[5]. This method has not been proposed to synthesize any kind of research literature, but only studies with quantitative results: "Meta-analysis is the analysis of empirical analyses" [42],

[4] For instance, the Journal of Industrial Ecology recently published a special issue on MA applied to LCA in 2012 (vol. 16) of which [15] is the editorial.

[5] Six meta-analyses were published at the same time in the field of economics [42,109–113]. See for instance Stanley [44] for a more comprehensive presentation. Standard references for technical aspects of meta-analysis are [93,108,114,115].

not theoretical ones. Applied to environmental evaluation methods, this methodology is thus relevant to review previously reported LCA studies outcomes.

Research syntheses aim at summarizing findings in such a way that clear and uncontroversial conclusions may be drawn from previous accumulated knowledge. Yet, estimates obtained with an LCA approach are characterized by large differences among study results. Even if different studies deal with a same issue, each one departs from previous literature by using different data sets, different methodological choices, etc. Research synthesis may thus appear as an especially difficult task when reviewing LCA literature. Compared to qualitative literature reviews, the original idea behind MA is to consider study results in the same way as any scientific phenomenon. Each reported result is viewed as an "observation" of a complex dataset, "no more comprehensible without statistical analysis than would hundreds of data points in one [LCA] study" [41]. MA may then be understood as a set of statistical techniques which allows to systematically summarize quantitative studies. It is a complementary method to narrative literature surveys that generally provide a more qualitative than quantitative analysis of estimate results.

Using econometrics methods, which are specific statistical techniques, MRA may be considered as a subset of MA. It allows to review and analyze previous results through a ceteris paribus reasoning [43]. By doing so, outcomes from many studies can be integrated and combined in such a way that comparison between their results becomes easier. MRA provides a quantitative summary of estimate results, such as mean estimates and confidence intervals of the quantitative results among studies. Compared to narrative literature surveys, the major contribution of MRA consists in modeling estimate result variations as a function of different factors. The use of specific econometrics methods allows then to statistically estimate and quantify their influence on study outcomes.

More formally, let the generic form of the linear regression model be the "original model" of the MRA equation:

$$Y = f(X) + \varepsilon = X\beta + \varepsilon \tag{1}$$

where Y is the $(I \times 1)$ dependent variable vector composed of the I reported estimates of the phenomenon of interest in the MA. For reasons that will be developed in Section 3.1.3, the reported estimates of a MA are named "*e-s*" estimates. These I estimates are drawn from J studies. Note it is generally stated $I \geq J$. If only one estimate per study is retained, then $J=I$. As usual, the term ε is a $(I \times 1)$ vector of a random disturbance. It is assumed that the sampling error is normally distributed with mean zero and variance $\sigma^2_{\varepsilon,i} : \varepsilon_i \sim N\left(0, \sigma^2_{\varepsilon,i}\right), \forall i = 1, \ldots, I$. X is the $(I \times K)$ matrix composed of the K−1 independent variables of this meta-model. The independent variables represent study characteristics that are supposed to have an influence on the systematic excess variation of Y. β is the $(K \times 1)$ vector of the coefficients of this meta-model. Once estimated, it gives a measure of the particular effects of each characteristic.

The following notational convention will apply in the remaining of this paper: let the first column of the $(I \times K)$ data matrix, $X_{(I,K)}$, be a column of 1 s and the others column vectors be the I observations of the K-1 independent variables:

$$\underset{(I,K)}{X} = \left(\underset{(I,1)}{C}, \underset{(I,1)}{X_1}, \ldots, \underset{(I,1)}{X_l}, \ldots, \underset{(I,1)}{X_{K-1}} \right) \tag{2}$$

$$\text{where } C_{(I,1)} = \begin{pmatrix} 1 \\ \vdots \\ 1 \\ \vdots \\ 1 \end{pmatrix} \text{ and } \underset{(I,1)}{X_l} = \begin{pmatrix} x_{l,1} \\ \vdots \\ x_{l,i} \\ \vdots \\ x_{l,I} \end{pmatrix}$$

Let us specify the $(I \times 1)$ vector of the coefficients, $\beta_{(I,1)}$, as follows:

$$\underset{(K,1)}{\beta} = \begin{pmatrix} \alpha \\ \beta_1 \\ \vdots \\ \beta_l \\ \vdots \\ \beta_{K-1} \end{pmatrix}$$

According to this notational convention, $x_{l,i}$ is the i-th observation of the l-th independent variable ($i = 1, \ldots, I$ and $I = 1, \ldots, K-1$). β_l is the coefficient of the vector of the I observations of the *l-th* independent variable, X_l, and α is the constant term in the model, also known as the intercept.

In MRA dealing with LCA studies, X could be stated as being composed of three kinds of variables. $X_{(I,K)} = (C_{(I,1)}, T_{(I,t)}, M_{(I,m)}, S_{(I,s)})$ where T, M and S are assumed to be $(I \times t)$, $(I \times m)$ and $(I \times s)$ vectors, respectively. T is composed of t variables related to technical characteristics of pathways assessed in the primary studies. In this MRA, it corresponds to biofuel characteristics such as the type of biomass feedstock, the type of technologies and associated yields, etc. The m variables of M refers to methodological assumptions reflecting researcher choices: for instance the type of LCA approach (A-LCA or C-LCA), the system boundaries, etc. Finally, the s variables of S correspond to the typology of the study under consideration such as the type of this study (peer reviewed or working paper for instance), the publication year or the geographical location of authors. Of course, the definitive specification of Eq. (1) depends on both the particular issue investigated (here, Global Warming impact indicator of advanced biofuels) and studies reviewed in the MRA[6].

3. Database of LCA results of GHG emissions for advanced biofuels

3.1. Construction and composition of the database

As mentioned before, the goal of this study is to explain the variations of LCA results for GHG emissions of advanced biofuels. Consequently, the variable of interest (so-called *e-s* or dependent variable) is the result for GHG emissions per MJ of biofuel calculated with an LCA approach. These estimates have been drawn from the study sample of this MA. One value for GHG emissions (i.e. the estimate) corresponds to one observation in our MA sample. As one study can contain several estimates, our database (i.e. our MA sample) can be composed of more than one observation per study ($I \geq J$), recall Eq.(1).

The inclusion of all estimates from a single study is a source of disagreement in the MA literature. Some authors believe that only one estimate should be included per study based either on the mean of the available estimates, or selected on the basis of expert judgment, while other authors advocate including all estimates as a method of boosting sample size (see Stanley [44] for a discussion on this issue). We choose to include all estimates from a single study for the following two reasons. First, the choice of a particular estimate is subjective, and when facing the same estimates, different researchers may undoubtedly make different choices. To maintain a position as neutral as possible, we considered all available explicit results in the study or which are easily inferred. Second, the core of MA is to summarize quantitative literature in a systematic way regardless of its quality. Hence, it would not be

[6] See Appendix B for a more technical presentation on the treatment of heteroskedasticity in MRA.

Table 1
List of categories and subcategories of variables included in the database.

Technical data	Methodological choices	Typology of the study
Type of biofuel	Type of LCA approach	Type of study
Type of biomass feedstock	System boundaries	Year of publication
Type of coproducts	Method for taking into account coproducts	Geographical location of authors
Type of technologies and associated yields	Carbon neutral	
Geographical location of the case study	Characterization method for impact assessment	
	Method for assessing N_2O emission from N input	
	Method for taking into account Land Use Change	
	Method for taking into account uncertainties	
	Number and type of environmental impact indicator assessed in the study	

relevant to select studies ex-ante regarding their quality since this choice would be arbitrary. The MRA literature proposes various ex-ante tests (such as statistical ones) that can lead to exclude some studies ex-post, or at least some of their estimates, from the database/MA sample.

3.1.1. *Selection and description of studies*

Before proceeding to a MA, the database of the MA has to be constituted. To do so, some common procedures exist in MA. Stanley [44] describes three steps to conduct a MA. First, primary studies having estimated a common quantitative effect are identified among published and unpublished literature. This set of studies is the material of the MA. Second, each article results and features are coded in a database. By doing so, studies are characterized in a way that allows them to be compared. Their findings, i.e. their estimates, become the observed values of the dependent and independent meta-variables. The *e-s* and potential factors which are supposed to have any influence on its variations are identified and summarized in a coded form: the explanatory variables of the matrix X. Third, the MRA can be conducted to estimate the magnitude of the quantitative effect under consideration and better understand variations in the reported estimates.

This section details the selection process of studies included in this MA. To obtain and analyze estimates for the GHG emissions of advanced biofuels, a large bibliographical research has been carried out to collect studies using an LCA approach. We have taken a census of both published articles and "grey literature", such as unpublished papers, conference papers, official reports. The existence of published articles presenting detailed literature reviews dealing with issues that are similar to ours has already been mentioned: [9–14]. These literature reviews were the starting point of the bibliographic research. Entries of their bibliographic references have been systematically reviewed. Then, to complete this first paper selection, a web-based keyword search - e.g. "LCA", "biofuel", "second generation biofuel", "third generation biofuel", "advanced biofuel", "cellulosic ethanol", "lignocellulosic ethanol", "synthetic diesel", "syndiesel", "BTL", "microalgae", "microalgae biodiesel", etc. - has been done on relevant literature databases (Science Direct, Web of Science, SciVerse, Springer Link, etc.) and web sites of major publishers of academic journals (Blackwell, Elsevier, Kluwer, Sage, Springer, Taylor Francis, and Wiley). The "grey literature" has been more particularly collected through Google and Google Scholar, Dissertation Abstracts, web sites of key academic institutions and authors and web sites of major environmental evaluation conferences.

To better insure the homogeneity of the sample, studies have to meet four selection criteria to be included in the sample of this MA: (i) only studies with primary results were included to avoid double counting (no literature reviews)[7] , (ii) only studies using an LCA approach were included [8], (iii) only LCA studies on the following liquid transportation fuels were included: lignocellulosic ethanol, FT diesel, microalgae HAO and FAME[9], (iv) only studies assessing Global warming impact indicator (i.e. GHG emissions) with "Well To Tank" (WTT) or "Well To Wheel" (WTW) boundaries[10]. The proxy used to measure the GHG emissions has to be the expressed (or easily convertible) in term of grams of CO_2 equivalent per MJ of biofuel.

Moreover, no a priori filter was used concerning the type of publication (published or unpublished papers) but the date and the English language. This MA focuses on studies conducted since 2002 (until mid 2011) since, to our knowledge, no advanced biofuels LCA studies were conducted before this date.

At the end of this selection process, the database contains 47 LCA studies [5,6,32,45–87] providing 593 estimates of life-cycle GHG emissions of advanced biofuels. Details of number of estimates by studies included in the sample are provided in Table 2 (see Table I.1 in the Supplementary data for details about selected studies).

3.1.2. *Choice and description of the meta-variables*

The object of this MA is twofold. First, this MA proposes a statistical summary of the role of different determinants for estimates of the *e-s*, i.e. the Global warming impact indicator for advanced biofuels in grams of CO_2eq per MJ. By identifying and measuring the influence of these determinants, one may obtain a more in-depth explanation of how advanced biofuel LCA GHG emission estimates change as these factors vary. Second, an important aspect of this article is to provide average estimates of the Global warming impact indicator for advanced biofuels.

The dependent (*e-s*) and independent variables (potential factors) of this MA are now detailed.

3.1.3. *The effect-size: the dependent variable*

As mentioned before, the variable of interest (*e-s* or dependent variable) is the result for GHG emissions per MJ of biofuel

[7] The MA literature distinguishes primary studies from secondary ones. Compared to the latter, the former presents original research results. Litterature reviews are the typical example of secondary studies. In order to avoid double counting, only results drawn from primary studies are included in a meta-database.

[8] Only studies following the ISO 14044 guidelines to conduct an LCA were included [21].

[9] Studies on other biomass derived fuels such as methanol, DME, ETBE, biogas, heat, power, CHP were not included for reasons already mentioned in the introduction of this paper.

[10] To be more precise, only the WTW studies with consumption of pure biofuels have been included. Studies containing aggregate results for fuel blends such as E10 (blend of 10% ethanol and 90% gasoline) were not included in the database. No study with a bi-functional unit was included.

Table 2
List of selected studies for the MA with a description of some of their characteristics.

Study	# of Obs	Year	e-s (mean in g CO_{2eq}/MJ)	Type of biofuel generation	Type of LCA approach	Uncertainty analysis? (method)[a]	LUC?	Type of Study (PR, OR, Dir., WP)[b]	Geographical location of authors
Bai et al. [45]	2	2010	27.36	G2 (Ethanol)	A-LCA	No	No	PR	Europe
Batan et al. [46]	14	2010	−55.43	G3	A-LCA	Yes (SA)	No	PR	North America
Campbell et al. [47]	6	2010	−9.42	G3	A-LCA	No	No	PR	Other
Cherubini et al. [48]	6	2011	41.07	G2 (Ethanol)	A-LCA	No	No	PR	Europe
Choudhury et al. [49]	3	2002	25.03	G2 (Ethanol & BtL)	A-LCA	Yes (MC)	No	WP	Europe
Chouinard-Dussault et al. [50]	4	2006	39.64	G2 (Ethanol)	A-LCA	No	Yes	WP	North America
Delucchi [51]	7	2010	−19.29	G2 (Ethanol)	A-LCA	No	Yes	PR	North America
Elsayed et al. [52]	1	2003	13.00	G2 (Ethanol)	A-LCA	No	No	WP	Europe
Fazio and Monti [53]	15	2011	16.80	G2 (Ethanol & BtL)	A-LCA	No	No	PR	Europe
González-García et al. [54]	8	2010	114.96	G2 (Ethanol)	A-LCA	Yes (SA)	No	PR	Europe
González-García et al. [55]	1	2010	35.39	G2 (Ethanol)	A-LCA	No	No	PR	Europe
Gonzalez-Garcia et al. [120]	1	2009	−9.99	G2 (Ethanol)	A-LCA	No	No	PR	Europe
Groode and Heywood [56]	4	2007	9.75	G2 (Ethanol)	A-LCA	Yes (MC)	No	WP	North America
Haase et al. [57]	2	2009	15.53	G2 (BtL)	A-LCA	No	No	WP	Europe
Hoefnagels et al. [58]	90	2010	12.94	G2 (Ethanol & BtL)	A-LCA	No	Yes	PR	Europe
Hsu et al. [59]	8	2010	41.89	G2 (Ethanol & BtL)	A-LCA	Yes (MC)	No	PR	North America
JEC [60]	6	2007	11.52	G2 (Ethanol & BtL)	A-LCA	Yes (MC)	No	OR	Europe
JEC [61]	6	2011	11.77	G2 (Ethanol & BtL)	A-LCA	Yes (MC)	No	OR	Europe
Jungbluth et al. [62]	9	2007	61.29	G2 (BtL)	A-LCA	Yes (SA)	No	OR	Europe
Jungbluth et al. [63]	22	2008	47.90	G2 (BtL)	A-LCA	No	No	OR	Europe
Kaufmann et al. [64]	25	2010	24.53	G2 (Ethanol)	A&C-LCA	Yes (SA)	No	PR	North America
Koponen et al. [65]	108	2009	43.85	G2 (Ethanol)	A-LCA	Yes (SA)	Yes	WP	Europe
Lardon et al. [66]	4	2009	94.00	G3	A-LCA	No	No	PR	Europe
Luo et al. [67]	9	2009	163.84	G2 (Ethanol)	A-LCA	No	No	PR	Europe
McKechnie et al. [68]	6	2011	−55.88	G2 (Ethanol)	A-LCA	No	No	PR	North America
Mehlin et al. [69]	2	2003	8.28	G2 (BtL)	A-LCA	Yes (SA)	No	WP	Europe
Mu et al. [70]	19	2010	−5.33	G2 (Ethanol & BtL)	A-LCA	Yes (SA)	No	PR	North America
Mullins et al. [71]	10	2010	41.10	G2 (Ethanol)	A-LCA	Yes (MC)	Yes	PR	North America
RED [5]	10	2009	12.80	G2 (Ethanol & BtL)	A-LCA	No	No	OR/Dir.	Europe
RFS2 [6]	12	2010	20.67	G2 & G3	C-LCA	Yes (MC)	Yes	Dir.	North America
Sander et al. [72]	1	2010	−18.40	G3	A-LCA	No	No	PR	North America
Schmitt et al. [73]	3	2011	49.62	G2 (Ethanol)	A-LCA	No	No	PR	North America
Sheehan et al. [74]	1	2004	−81.28	G2 (Ethanol)	A-LCA	No	Yes	PR	North America
Spatari et al. [75]	2	2005	18.94	G2 (Ethanol)	A-LCA	No	Yes	PR	North America
Spatari et al. [76]	34	2009	−2.69	G2 (Ethanol)	A-LCA	Yes (MC & SA)	Yes	PR	North America
Spatari et al. [77]	6	2010	−7.93	G2 (Ethanol)	A-LCA	Yes (MC)	Yes	PR	North America
Stephenson et al. [78]	17	2010	12.12	G2 (Ethanol)	A-LCA	Yes (SA)	No	PR	Europe
Stephenson et al. [79]	31	2010	201.15	G3	A-LCA	Yes (SA)	No	PR	Europe
Stichnothe and Azapagic [80]	18	2009	33.98	G2 (BtL)	A-LCA	Yes (SA)	No	PR	Europe
Stratton et al. [81]	23	2010	24.60	G2 & G3	A-LCA	Yes (SA)	Yes	WP	North America
Van Vliet et al. [82]	5	2009	−15.78	G2 (BtL)	A-LCA	No	No	PR	Europe
Vera-morales and Schäfer [83]	4	2009	55.75	G3	A-LCA	No	No	WP	Europe
Wang et al. [84]	3	2010	13.79	G2 (Ethanol)	A-LCA	No	Yes	PR	North America
Wang et al. [85]	3	2011	8.00	G2 (Ethanol)	A-LCA	No	No	PR	North America
Wang et al. [85]	15	2011	57.50	G2 (Ethanol)	A-LCA	No	Yes	PR	Europe
Wu et al. [86]	5	2005	14.72	G2 (Ethanol & BtL)	A-LCA	No	No	OR	North America
Xie et al. [87]	2	2011	−59.24	G2 (BtL)	A-LCA	Yes (MC)	No	PR	North America
Number of studies	47								
Number of observations	593								
Mean and repartition (weighted by observations)		2009	34.45	G2 (87%) of which BtL (26%) and ethanol (61%), G3 (13%)	A-LCA (97%), C-LCA (3%)	MC (10%), SA (38%), no uncertainty analysis (52%)	LUC (51%), no LUC (49%)	PR (65%), OR (12%), Dir. (4%), WP (19%)	North America (45%), Europe (53%), Other (2%)
Mean and repartition (weighted by studies)	13	2009	23.07	G2 (87%) of which BtL (38%) and ethanol (70%), G3 (17%)	A-LCA (98%), C-LCA (4%)	MC (21%), SA (26%), no uncertainty analysis (53%)	LUC (28%), no LUC (72%)	PR (65%), OR (12%), Dir. (4%), WP (19%)	North America (45%), Europe (53%), Other (2%)
Median (weighted by studies)	6	2010	15.53						

[a] MC=Monte Carlo analysis, SA=sensitivity analysis.
[b] PR=Peer review, OR=Official Report, Dir.=legislative text (Directive or Standard), WP=Working Paper.

calculated with an LCA approach. Those estimates drawn from different studies, i.e. the observations of our MA sample, may be expressed in different units of measure. These values need to be converted in a way that allows them to be combined to constitute the meta-dependent variable. The transformation of the dependent variable observations into a unique metric measure is a

common procedure in MA studies. This step is called the *e-s* calculation and is central to MA literature. Indeed, it is this conversion of the dependent variable in a standard measure, the *e-s*, that allows to compare previous results and to investigate their determinants. In our sample, most of the studies present the GHG emissions, in grams of CO_2 equivalent, as a midpoint impact category using IPCC's characterization factors. Some other studies present only inventory data on GHG emissions so these results had to be converted into grams of CO_2 equivalent. We used the latest IPCC characterization factors [88] for these conversion steps. It was not possible to harmonize all of the observations by using the IPCC's 2007 characterization factors because inventory data (individual GHG emissions) were not always available. It has been shown, however, that the calculation method for global warming impact has an insignificant influence in LCA results [37,89].

Still, there is another step in the calculation of the e-s since the LCA results are not always presented for the same functional unit. Typical functional units in biofuel LCA studies are a unit of fuel produced (liter, kg, MJ, etc.) or the service rendered by the biofuel (dislocation of a vehicle for a certain distance expressed in km, miles, etc.). Some other studies present their results using other less conventional functional units such as the surface of arable land used. All of these choices depend on the initial goals of the study.

We choose to convert the GHG emission values in our database into a common functional unit, a MJ of fuel produced since this is the unit used in the RED (the RFS2 also presents results for biofuel energy content, in Btu). For a given study, we apply conversion factors using the provided information in the study for lower heating values (LHV), densities, engine fuel consumption, etc. Whenever these values did not appear in a study, information from a well-documented study was used [90]. Some studies had to be discarded because results were presented for a functional unit that could not be converted into a MJ (e.g. Melamu et al. [91]. is a C-LCA study where the results are presented for a multi-functional unit, involving fuel and electricity production).

Lastly, a standard error is associated to every observation so that our sample can be treated for heteroskedasticity. As mentioned before, there are mainly two ways to treat uncertainty in LCA (and consequently estimate standard errors): Monte-Carlo analysis and sensitivity analysis. The standard error could be directly inserted in the database only for the observations from studies performing Monte-Carlo analysis. We calculated a standard error from the e-s variance of each sensitivity analysis performed (one study can present the sensitivity of LCA results for variations of more than one parameter, each performed separately). For the studies that did not assess the uncertainty of their results, we calculated the standard error based on all the available observations for a same type of fuel.

3.1.4. *The potential factors: the independent variables*

There are no guidelines concerning exactly which variables, potentially influencing LCA results, have to be included in a MA independent variable set. Like any other scientific investigation, this choice is determined by the available data [92], LCA practitioner knowledge (see Section 2.1) and the specificities of each technology (see Appendix A). Some non-intuitive variables are also included in the database. In addition, some study characteristics (country, year of publication, etc.) were included to account for potential publication biases.

Primary studies highlight different determinants of advanced biofuel GHG emission estimates whereas surveys offer a more in-depth discussion on their likely influences. According to the introduction of this section, three categories of potential determinants of GHG emission estimates are kept: technical data,

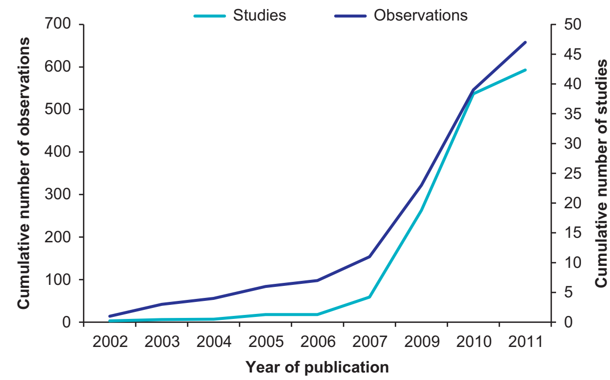


Fig. 2. Cumulative number of studies and observations per year of publication.

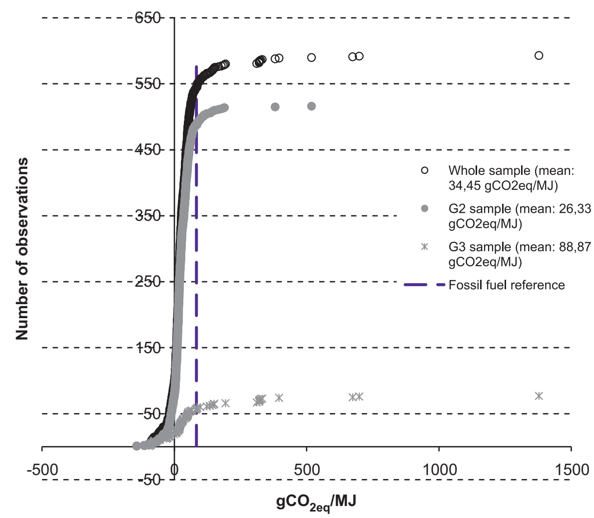


Fig. 3. Dispersion of LCA GHG emission results included in the database for the different types of biofuel.

methodological choices of authors and typology of the study under consideration. The latter variables are more particularly based on typical variables employed in previous MA.

The three categories of explanatory variables are broken down further as follows. Each category could be divided into subcategories (see Table 1). Those subcategories could gather from 2 to 18 variables. All variables are encoded either as binary—a.k.a. dummy or qualitative—variables or as quantitative variables. At present, more than 80 variables are available in the database.

A brief description of all subcategories for all categories follows (see Table I.2 in the Supplementary Data for a comprehensive variable description and their respective names):

3.1.4.1. *Technical data.*

The type of biofuel (Biomass To Liquid, Ethanol, Fatty Acid Methyl Ester or Hydrotreated Algal Oil) as well as the biofuel generation (G2 biofuel for BtL and Ethanol; G3 biofuel for FAME and HAO) are set as variables.

In the “type of biomass feedstock” category, due to the variety of feedstock used for biofuel production in our sample, we created groups for biomass having similar characteristics (e.g. poplar and eucalyptus are coded as farmed wood, corn stover and wheat straw are coded as agricultural residues, etc.). An additional variable was created in order to test the difference of using cultivated resources (energy crops and farmed wood) and waste/residues as feedstock (biomass from agricultural or forestry residues) on LCA results.

In the “type of technologies and associated yields” category, all different types of processes for biomass pretreatment and for

Table 3
Statistical description of GHG emission results included in the database for the different types of biofuel and for the different geographical location of authors.

Biofuel generation	Location of authors	# of Obs.	(%)	Median[a]	Mean[a] [Confidence Interval]	Standard deviation	Extrema[a] Min	Percentiles[a] Max	5th	95th
G3 & G2	All	593		21.60	34.45 [27.26;41.64]	89.34	−142.18	1377.90	−37.08	116.65
	North America	198	(33%)	12.61	4.72 [−1.24;10.68]	42.78	−142.18	193.20	−79.66	55.40
	Europe	401	(68%)	26.05	48.47 [38.53;58.41]	101.54	−88.36	1377.90	2.44	144.68
G3	All	77		31.00	88.87 [41.55;136.19]	211.85	−96.47	1377.90	−85.00	332.20
	North America	38	(49%)	17.99	0.22 [−21.62;22.05]	68.67	−96.47	193.20	−89.89	134.98
	Europe	45	(58%)	61.86	150.63 [76.58;224.68]	253.44	−30.97	1377.90	8.69	676.39
G2	All	516		20.50	26.33 [22.43;30.23]	45.20	−142.18	518.40	−24.00	85.80
	North America	160	(31%)	12.41	5.79 [0.51;11.08]	34.12	−142.18	71.00	−60.07	49.47
	Europe	356	(69%)	24.25	35.56 [30.72;40.39]	46.55	−88.36	518.40	1.00	100.76
G2-BtL	All	155		14.50	19.04 [13.41;24.68]	35.78	−142.18	189.00	−18.50	69.05
	North America	36	(23%)	6.10	−1.55 [−12.67;9.57]	34.05	−142.18	47.61	−54.08	32.15
	Europe	119	(77%)	15.80	25.28 [19.16;31.39]	34.03	−88.36	189.00	2.11	85.76
G2-Ethanol	All	361		24.30	29.45 [24.46;34.45]	48.39	−113.60	518.40	−25.56	89.78
	North America	124	(34%)	15.39	7.93 [1.95;13.91]	33.97	−113.60	71.00	−61.12	49.99
	Europe	237	(66%)	30.87	40.72 [34.22;47.21]	50.99	−42.00	518.40	1.00	104.55

[a] Expressed in g CO_{2eq}/MJ.

conversion into fuel that we found in the literature were set as variables for BtL and Ethanol technologies. The "Mass yield provided" variable indicates if a value for a mass yield of the biofuel process unit is available in the study (this can be seen as a quality indicator for a given study) and the "Value of mass yield" indicates this value only for G2 biofuels. For G3 biofuels, we choose the daily productivity and the oil content of microalgae as quantitative variables since they have been often identified in the literature as the most influencing factors for life cycle GHG emissions of G3 biofuels. In addition, the fact of growing microalgae in open ponds or photobioreactors is set as a variable.

3.1.4.2. Methodological choices. All classical methodological choices for LCA are set as variables. We differentiate LCA studies with an attributional approach from LCA studies with a consequential approach (see Section 2.1)

Some hypothesis relative to system boundaries are set as variables: we distinguish WTT from WTW studies and the inclusion, or not, of infrastructures within the system boundaries is also taken into account.

As highlighted in Section 2.1, the methods used to account for coproducts can have a great influence in biofuel LCAs. Therefore they were also set as independent variables. We classify the observations as either using an allocation method (based on energetic, mass content, market value, etc.) or system expansion method. Some studies mix both methods, which we call hybrid method.

The carbon-neutrality hypothesis is very common in G1 and G2 biofuel studies. However, this hypothesis is not straightforward for studies involving microalgae since they do not always capture CO_2 directly from the atmosphere. CO_2, from flue gas for example, is generally fed into the system. Therefore, the carbon-neutrality hypothesis is set as an independent variable for G3 biofuels.

In order to study the influence of the choice of a characterization method for impact assessment, we make a distinction between studies that take into account 3 GHGs (CO_2, CH_4, N_2O) and studies that take into account more than 3 GHGs.

As also mentioned in Section 2.1, N_2O emissions from the field play an important role in the GHG emissions of biofuel lifecycles. The use of IPCC's method [30] or other more complex methods for estimating these emissions are set as independent variables.

Studies that take into account direct, indirect or both Land Use Changes for GHG emission calculation are also identified. The method for taking into account uncertainties is identified in each study: uncertainty analysis could be conducted by a Monte Carlo analysis or by a sensitivity analysis on specific factors (ceteris paribus) or no uncertainty analysis (recall Section 2.1). We also try to identify if the fact that a study assess other environmental impacts than GHG emissions could influence the GHG emission results. So the number and type of environmental impact indicators assessed in the study is controlled.

3.1.4.3. Study typology. Other aspects than technical data or methodological choices are included in the database. The type of study is identified: it can be classified as peer reviewed literature, official report, legislative text (Directive or Standard) or working paper. The year of publication as well as the geographical location of the authors is also included in the database.

3.2. Description of the database

This section deals with the statistical description of the database, which covers a large portion of studies that explicitly used LCA to evaluate environmental impacts of advanced biofuels. Finally, 47 LCA studies have been selected representing 593 observations of GHG emission results representing an average of 13 observations per study (see Table 2). Subsequently, this database is used to perform the MRA (see Section 4).

As displayed in Table 2, 87% of the studies in the database assess G2 biofuels (38% of studies assessing BtL and 70% Ethanol) and 17% of the studies assess G3 biofuels. Thus, among the 593 observations included in the database, those for G3 biofuels represent 13%. The other observations correspond to G2 biofuels of which 30% are for BtL and 70% are for Ethanol. Most of the studies adopt an attributional LCA approach; only 3% of the observations are calculated with a consequential LCA approach. Half of the studies do not perform an uncertainty analysis on their results. Among studies that include an uncertainty analysis, 44% perform a Monte Carlo analysis. Only 28% of studies included in the database take into account LUC (and only 4% address Indirect LUC), representing 51% of the observations. Observations extracted from peer reviewed literature represent 61% of observations (65% of studies), from the official reports 9% (12% of studies), from regulatory texts 3% (4% of studies), and from working papers 25% (19% of studies).

Furthermore, we can observe in Fig. 2 that the number of studies assessing GHG emissions of advanced biofuels increased sharply from 2007. This phenomenon could be linked with the

publication of legislative texts in the EU and the US regarding mandatory GHG emission savings thresholds for biofuels (respectively RED in 2009 and RFS2 in 2010).

3.2.1. Observations per type of biofuels

As depicted in Fig. 3 (see also Table 3), the mean value in the literature for G3 biofuels GHG emissions is quite similar to GHG emissions for the fossil fuel reference as defined in EU and US regulations—respectively 83.8 g CO_2eq/MJ (same reference for gasoline and diesel) and 92.5 g CO_2eq/MJ (mean of US gasoline and diesel references). GHG emissions mean value for G2 biofuels indicates that they can induce a GHG emission reduction compared to the fossil fuel reference from 69% to 72% (depending on the fossil fuel reference chosen). Therefore, from a statistical point of view, G3 biofuels seem to emit more GHG emissions during their life cycle than G2 biofuels. In the same way, GHG emissions mean for BtL is lower than for Ethanol (GHG emission savings compared to fossil fuel reference from 77% to 79% for BtL and from 65% to 68% for Ethanol).

The range of GHG emission results for G3 biofuels is very wide compared to the one for G2 biofuels as illustrated by their standard deviations (see Table 3). Hence, G3 biofuels could emit 20 times more GHGs than the fossil fuel reference whereas G2 biofuels could emit from 4 to 9 times more by considering the highest values of the literature results. Conversely, the lowest results are negative and quite similar for G2 and G3 biofuels.

Even though LCA results are inconclusive regarding GHG emission performances of advanced biofuels due to their wide range of variation, some trends can be indentified: on average, GHG emissions for G3 biofuels are higher than for G2 biofuels and GHG emissions for Ethanol are higher than for BtL. Thus, the type of biofuel seems to be an explanatory variable for the differences between the GHG emission results for advanced biofuels.

3.2.2. Observations per regions

We make the distinction between the geographical location of the authors (affiliation of the first author) and the geographical location of the cases studies (i.e. geographical location of inventory data). Regarding the geographical location of the authors, 45% of studies are from North American (NA) authors (including US and Canada) and 53% are from European authors (including EU countries and Switzerland), representing 32% and 67% of the observations respectively (see Table I.3 in the Supplementary Data). The other study is from Australian authors [47]. For G3 biofuels, 42% of observations are from NA authors, 51% from European authors and 7% from Australian authors. For BtL, 23% of observations are from NA authors and 77% from European authors. For Ethanol, 34% of observations are from NA authors and 66% from European authors. In most of the studies, the geographical location of the authors fits with the geographical location of the assessed pathways. Only 3% of the observations do not match ([67] and some observations of [58]). Therefore, we focus only on the geographical location of the authors as a measure of the potential influence of geographical location on GHG emissions.

On average for all types of biofuel, GHG emission results from NA authors seem to be lower than from European authors with a gap that could be significant as illustrated in Table 3 (e.g. from 0.22 g CO_2eq/MJ for NA to 150.63 g CO_2eq/MJ for Europe for G3 biofuels). Hence, it seems that the geographical location of the authors can have an influence on the GHG emission variability observed for advanced biofuels.

Figs. 4 and 5 present the dispersion of GHG emission results included in the database for the different types of biofuel and for the different geographical locations. These results are also

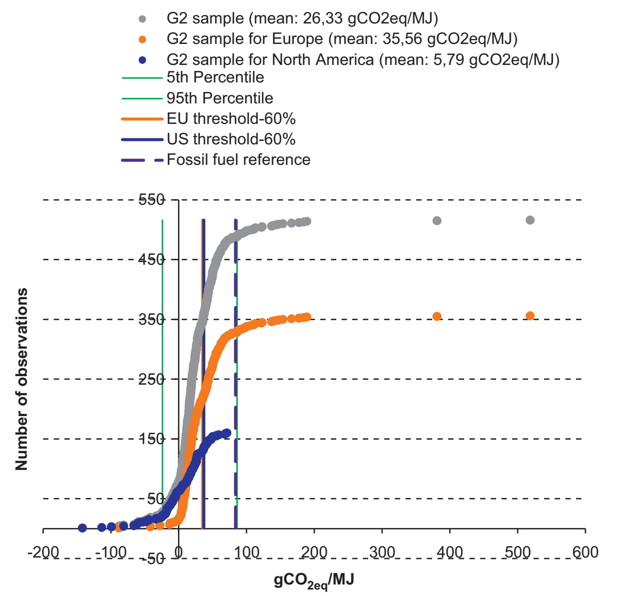


Fig. 4. Dispersion of LCA GHG emission results included in the database for G2 biofuels and for the different geographical location.

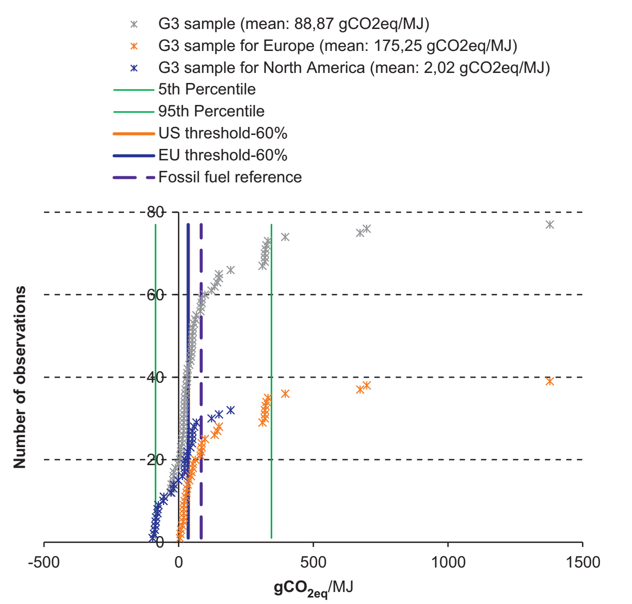


Fig. 5. Dispersion of LCA GHG emission results included in the database for G3 biofuels and for the different geographical location.

compared with their respective GHG emission minimum threshold depending on their geographical location.

As already mentioned, the RED and RFS2 set minimum GHG emission savings for biofuels. Their more restrictive savings are set to 60% compared to their corresponding fossil fuel reference (fossil fuel references are slightly different). According to Fig. 3, 82% of GHG emission results from NA are compliant with their more restrictive GHG emission minimum threshold whereas only 59% from Europe are compliant with their corresponding threshold. At this stage of the analysis, we do not have objective reasons explaining this systematic difference between NA and EU estimates. It may come from the use of a different set of technical variables, for instance, but it may also reveal the existence of a potential publication bias in the literature.

In conclusion, this section based on descriptive statistics allows the formulation of some collective insights from literature LCA results about factors that could influence GHG emission results for advanced biofuels. The type of biofuels (G2 vs. G3 biofuel, BtL vs. Ethanol) and the geographical location (North America vs. Europe) seem to have an influence on the variability of GHG emission results for advanced biofuels. However it is not possible to be more conclusive and accurate with the descriptive statistics presented in this section. Descriptive statistics and inspection of graphics are very useful and often relevant but remain always vulnerable to subjective interpretation. Thus, more objective statistical tests are needed, as those that could be done with MRA. By using specific econometrics methods, we believe that a MRA should allow the (i) confirmation of the insights previously identified and (ii) to go further in the explanation of the variability by identifying and quantifying the main variation factors.

Let us now develop the MRA based on these LCA studies.

4. Meta-regression analysis

Compared to narrative literature reviews, the MRA methodology allows us (i) to statistically identify main drivers of the *e-s* variability and (ii) to estimate both the direction and the magnitude of their respective effects across primary studies under consideration. The logic of MRA is illustrated here by applying this methodology to LCA literature evaluating GHG emissions of advanced biofuels. We first present the MRA model and its results for various G2 and G3 biofuel sub-samples. Second, we use the technique of *benefits transfer using meta-regression models* to propose a first attempt of harmonization of these LCA results.

4.1. The meta-regression model

Simply stated, to review a specific environmental evaluation literature, one must summarize its previous results already published on the issue under consideration.

It may be convenient to refer to a single observation in Eq. (1). Then, Eq. (1) may be rewritten as follows:

$$y_i = \alpha + \beta_1 x_{1,i} + \beta_2 x_{2,i} + \ldots + \beta_l x_{l,i} + \ldots + \beta_{K-1} x_{K-1,i} + \varepsilon_i, \forall i = 1, \ldots, I = \alpha + \sum_{l}^{K-1} \beta_l x_{l,i} + \varepsilon_i, \forall i = 1, \ldots, I = \underset{(1,K)(K,1)}{X_i' \; \beta} + \varepsilon_i, \forall i = 1, \ldots, I \quad (3)$$

where

$$\underset{(I,K)}{X} = \begin{pmatrix} \underset{(1,K)}{X_1'} \\ \underset{(1,K)}{X_2'} \\ \vdots \\ \underset{(1,K)}{X_i'} \\ \vdots \\ \underset{(1,K)}{X_I'} \end{pmatrix} \quad \text{and} \quad \underset{(1,K)}{X_i'} = (1, x_{1,i}, \ldots, x_{l,i}, \ldots, x_{K-1,i}), \ \forall i = 1, \ldots, I$$

We consider I advanced biofuels GHG emission estimates, the *e-s*, indexed by $i = (1, \ldots, I)$ and assume that the "*true*" *e-s* value for a given estimate is given by[11]:

$$y_i = \alpha + \underset{(1,K)(K,1)}{X_i' \; \beta} + \mu_i, \forall i = 1, \ldots, I \quad (8)$$

where y_i is the true *e-s*, α is a common factor, X_i' is a vector that measures characteristics of the biofuel case study and of the study under consideration, β is a vector of parameters to be estimated, and μ_i is normally distributed with mean zero and variance $\tau^2_{\mu,i} : \mu_i \sim N(0, \tau^2_{\mu,i})$

The "*true*" *e-s* value, y_i, is not observed. Instead, each study provides an estimated *e-s*, y_i, so that:

$$y_i = y_i + \varepsilon_i = \alpha + \underset{(1,K)(K,1)}{X_i' \; \beta} + \mu_i + \varepsilon_i, \forall i = 1, \ldots, I \quad (9)$$

where ε_i is an error term that is normally distributed with mean zero and variance $\sigma^2_{\varepsilon,i} : \varepsilon_i \sim N(0, \sigma^2_{\varepsilon,i}), \forall i = 1, \ldots, I$

Thus we allow the "*true*" *e-s* and the precision of the estimated *e-s*, $\sigma^2_{\varepsilon,i}$, to vary across estimates. The term $\sigma^2_{\varepsilon,i}$ is known as the within-variance and varies from study to study. As already mentioned, it is usually taken as given and derived from the original estimate.

Any remaining heterogeneity between estimates is either explainable by the observable differences modeled through the moderator variables contained in X_i' or is random and normally distributed with mean zero and variance $\tau^2_{\mu,i}$, the between-variance.

If $\tau^2_{\mu,i} = 0$, the model is referred to the fixed-effects model, and it is assumed that all heterogeneity in the "*true*" *e-s* can be explained by differences in study characteristics. If the between-variance is not equal to zero, the model is a random effects model (REM), which is usually referred to as a "mixed-effects" model because it contains observable "fixed" characteristics in X_i' as well as a random unobservable component with mean zero and variance $\tau^2_{\mu,i}$. The unknown variance can be estimated by an iterative (restricted) maximum likelihood process or, alternatively, using the empirical Bayes method, or a non-iterative moment estimator.

Note that the meaning of the adjectives "fixed" and "random" in the MA literature is different from the usual interpretation for panel data models in standard econometrics, because they refer to assumptions about the underlying population *e-s* [93]. In standard econometric terms, the fixed-effects meta-estimator is equivalent to the weighted least squares (WLS) estimator using the estimated variances (derived in the primary studies) as weights and re-scaling the standard errors of the meta-regression by means of the square root of the residual variance. The random effects estimator is akin to a random coefficient model in which the within- and between-study variances are used as weights [94].[12]

4.2. Meta-regression analysis results

Since the studies in the primary literature may use different data sets and different ways of modeling, we have good reasons to suspect that our sample is heteroskedastic.

A common approach is to use White's Heteroskedastic-Consistent Covariance Matrix (HCCM). This estimator simultaneously corrects for heteroskedasticity and cluster autocorrelation, and hence accounts for the multiple data setup by allowing different variances and non-zero covariances for clusters of measurements from the same study. As highlighted by [19], the White estimator [95] is *arguably rather restrictive assuming that all differences across observations and studies are observable and can entirely explain the empirical heterogeneity*. In addition, the White estimator does not fully exploit all available information because it estimates the variance rather than taking it as given or recoverable from the primary studies.

The latter can be remedied by using the fixed-effects meta-estimator that we already presented. As explained above, $\sigma^2_{\varepsilon,i}$ is a sample estimate of the standard deviation of the meta-regression errors. When this kind of measure of the heteroskedasticity

[11] The following presentation is partly inspired from Ready [117].

[12] Thompson and Sharp [118] provide an overview of various estimators that allow for random-effects variation.

Table 4
Results of MRA for the econometric samples Whole and G2 biofuels.

Samples Model	Whole 1aAll	Whole 2aAll	G2 1aG2	G2 2aG2	G2 1bG2	G2 2bG2
Constant	76.27*** (13.64)	271.74*** (23.66)	20.32*** (3.43)	27.24*** (4.72)	21.14*** (3.61)	28.43*** (5.04)
Technical data						
gen_3 (ref for Whole)						
etha	−41.39*** (13.14)	−220.92*** (23.77)	5.84*** (1.91)	–	5.83*** (1.81)	–
btl (ref for G2)	−52.12*** (13.36)	−215.57*** (22.59)				
mat_cult					−9.47*** (2.21)	−11.56*** (3.02)
mat_cultxdluc			−7.94*** (2.46)	−13.67*** (3.23)		
Methodological choices						
lca_att (ref)						
lca_cons			−33.66*** (4.79)	−40.41*** (8.63)	−34.04*** (4.9)	−39.09*** (8.37)
copval_alloc			8.96*** (1.91)	8.82** (3.62)	8*** (1.94)	6.99* (3.84)
copval_systexp (ref)						
copval_hyb			5.25** (2.38)	–	5.41** (2.69)	–
luc_dir					–	–
luc_indir			29.97*** (6.32)	39.78*** (7.27)	29.62*** (6.34)	36.54*** (7.2)
uncer_MC			8.03** (3.45)	16.68*** (4.61)	8.04** (3.41)	17.25*** (4.58)
uncer_SA			7.78*** (2.4)	7.08* (3.63)	7.32*** (2.39)	6.69** (3.39)
uncer_ref (ref)						
impcat_nev			9.26*** (2.99)	–	7.71*** (2.71)	–
impcat_nrc			−15.01*** (2.36)	−7.31** (3.41)	−12.65*** (2.54)	–
impcat_other			–	–	0.84* (0.49)	–
impcat_gwponly (ref)						
Typology of the study						
zlab_us	−24.6*** (3.97)	−190.58*** (25.05)	−8.32*** (2.1)	−18.58*** (3.66)	−8.66*** (2.12)	−19.73*** (3.2)
zlab_eu (ref)						
zlab_other	−85.69*** (15.6)	−281.16*** (24.85)				
Model information						
N	533	533	464	464	464	464
Mean dep. Var.	28.64	17.62	24.15	25.04	24.15	25.01
Adj. R-squ.	16.30%	68.76%	37.26%	30.95%	38.33%	30.94%
Log-Likelihood	−2727.20	−3068.04	−1976.89	−2044.50	−1972.36	−2044.03
F-stat. (*P*. value)	18.93 (0,0000)	32.82 (0,0000)				
Skewness (*P*. value)	61.27 (0.0000)		24.57 (0.017)		23.56 (0.0354)	
Kurtosis (*P*. value)	8.75 (0.0031)		1.6 (0.2062)		3.06 (0.0801)	
AIC	5464.39	6146.08	3977.78	4113.00	3970.72	4114.05
BIC	5485.79	6167.47	4027.45	4162.68	4024.54	4167.87
Wald Test (P. value) for etha=btl	26.29 (0.0000)	0.22 (0.6409)				
Procedure	OLS (White's HCCM)	WLS	OLS (White's HCCM)	WLS	OLS (White's HCCM)	WLS

is available, then Weighted Least Squares (WLS) becomes the obvious method to obtain efficient estimates of Eq. (9).

We start out by presenting the results obtained for the "*whole*" sample, which includes all the G2 and G3 biofuel studies included in the meta-database. Recall that our meta-database includes variables representing (i) technical data/characteristics, (ii) author's methodological choices and (iii) typology of the study under consideration. As technical data are specific to each type of biofuel, it is not possible to include this set of variables in the "*whole*" sample in order to test and quantify their respective influence. In order to capture characteristics of each biofuel generation and the type of fuel analyzed, one needs to break the "*whole*" sample into these respective sub-samples. In the subsequent sections we present the results for smaller samples named as follows: "*G3*", "*G2*", "*G2-BtL*" and "*G2-Ethanol*". Hence, the "*wholeº* sample corresponds to the merge of our "*G3*" and "*G2*" samples. Note that the "*G3*" and "*G2*" samples have been cut to 90% in order to exclude outliers which may have spurious influence on econometric estimates, as it is usually done in applied econometrics. So defined, the "*G2*" sample contains 464 observations (321 for Ethanol and 143 for BtL) and the "*G3*" sample contains 69 observations. (see Figs. 4 and 5 for a visual representation of "*G2*" and "*G3*" samples outliers). "*G2-BtL*" and "*G2-Ethanol*" sub-samples are a subset of the "*G2*" sample.

Results of Eq. (9) are presented in Table 4 for the "*whole*" and "*G2*" samples. Tables 5, 6 and 7 provide results for the "*G2-Ethanol*", "*G2-BtL*" and "*G3*" sub-samples respectively. For each model, results are systematically reported for two different corrections for heteroskedasticity: the first estimator uses the White's Heteroskedastic-Consistent Covariance Matrix (HCCM) (as denoted by the number 1 in columns) and the second one uses Weighted Least Squares (WLS) using inverse standard error weights (as denoted by the number 2 in columns)[13].

Unless it is indicated, all regression results are presented in reduced form. These models were chosen by the *general to specific* approach to econometrics modeling. As usual, "***", "**" and "*" respectively indicate 1%, 5% and 10% significance levels and standard errors of the coefficient estimates are reported in brackets. In each column, "–" means that the variable under consideration has been first included but finally removed from the reduced form because its coefficient estimate was not statistically significant at the 10% significance levels. Regarding model information, *N* and *Mean dep. Var* indicate respectively the number of observations used to perform each regression and the corresponding

[13] Each regression has been performed thanks to the STATA econometric software.

Table 5
Results of MRA for the econometric samples G2-Ethanol biofuels.

Samples Model	Ethanol 1aEtha	Ethanol 2aEtha	Ethanol 1bEtha	Ethanol 2bEtha	Ethanol 1cEtha	Ethanol 2cEtha
Constant	−5.88 (14.8)	31.9 (31.58)	46.39*** (8.33)	37.85*** (13.52)	32.11*** (2.64)	34.07*** (3.95)
Technical data						
mat_cultxdluc	−7.18** (3.26)	−10.1** (4.85)	−6.91** (3.28)	−9.92** (4.79)	–	–
g2_mass_yield			−73.57** (34.79)	–		
g2_mass_yield_sq						
g2_mass_yield_ln	−23.22** (9.3)	–				
Methodological choices						
lca_att (ref)						
lca_cons	−38.24*** (5.53)	−40.12*** (11.37)	−38.43*** (5.46)	−40.17*** (11.33)	−40.24*** (4.72)	−40.72*** (9.9)
luc_indir	29.01*** (7.49)	35.89*** (8.59)	27.48*** (7.41)	35.07*** (8.43)	19.85*** (7.29)	31.7*** (7.78)
uncer_MC	9.51** (4.06)	20.32*** (5.83)	9.66** (4.11)	20.43*** (5.82)	10.06** (4.23)	18.67*** (5.5)
uncer_SA	12.5*** (3.68)	12.28** (5.71)	11.53*** (3.59)	11.99** (5.57)	12.09*** (2.72)	14.01*** (4.08)
uncer_ref (ref)						
impcat_nev	12.34*** (3.99)	–	11.05*** (3.99)	–	9.51** (3.94)	11.15* (6.49)
impcat_nrc	−17.09*** (3.19)	−14.16*** (5)	−17.24*** (3.17)	−14.18*** (5)	−22.53*** (2.51)	−20.14*** (4.08)
impcat_other	–	–	–	–	−1.09* (0.61)	–
impcat_gwponly (ref)						
Typology of the study						
zlab_us	−7.29** (3.56)	−28.56*** (7.44)	−8.22** (3.56)	−29.23*** (7.35)	−11.26*** (2.37)	−23.84*** (4.74)
zlab_eu (ref)						
Model information						
N	209	209	209	209	321	321
Mean dep. Var.	19.70	19.14	19.70	18.96	26.61	26.61
Adj. R-squ.	31.21%	36.31%	30.32%	36.24%	40.13%	37.82%
Log-Likelihood	−884.15	−919.42	−885.49	−919.54	−1364.10	−1412.35
F-stat. (*P*. value)	12.61 (0,0000)	10.91 (0,0000)	12.55 (0,0000)	10.79 (0,0000)	33.94 (0,0000)	27.19 (0,0000)
Skewness (*P*. value)	18.14 (0.0527)		16.84 (0.078)		18.16 (0.0333)	
Kurtosis (*P*. value)	0.32 (0.5729)		0.23 (0.6351)			
AIC	1790.30	1860.83	1792.98	1861.08	2748.21	2844.70
BIC	1827.07	1897.60	1829.75	1897.85	2785.92	2882.42
LR test (*P*. value) Nested model: model (c)	959.9 (0,0000)	985.87 (0,0000)	957.22 (0,0000)	985.62 (0,0000)		
Procedure	OLS (White's HCCM)	WLS	OLS (White's HCCM)	WLS	OLS (White's HCCM)	WLS

mean of the dependent variable, i.e. the mean *e-s* expressed in g CO_2eq/MJ of biofuel.

In all tables, the quality of regressions is checked through the following diagnostic tests. Given that the simple R-squared statistic is sensitive to the number of variables included, only the adjusted R-squared is reported (*Adj. R-squ.*). The overall fit of the regression model is assessed by the logarithm of the Likelihood (*Log-Likelihood*) and the standard Fisher test, which tests for joint significance. The statistic of the latter test (*F-stat.*) and the corresponding measure of its statistical probability (*P.value*) are systematically reported. The null hypothesis of this test is all coefficients but the constant one is equal to zero. Two additional diagnostic tests for the quality of the regressions (and their *P. values*) are also reported: the Skewness's asymmetric test (*Skewness*) and the Kurtosis's normality test (*Kurtosis*) of residuals. They respectively correspond to a test of skewness and nonnormal kurtosis compared with the null hypothesis of symmetry (the skewness coefficient is zero for symmetrically distributed data) and kurtosis coefficient of 3. The normality tests examine the normality of the residuals – nonnormal residuals invalidate hypothesis tests on individual variables as these tests assume their normality. Therefore, this is an important consideration. All Tables also report the following two information criteria: the Akaike's Information Criterion (*AIC*) and the Schwarz's Bayesian Information Criterion (*BIC*). These two standard measures are used to allow (non-nested) model comparisons. Smaller *AIC* and *BIC* are preferred, because higher *Log-Likelihood* is preferred. Finally, in order to test and hence statistically confirm the importance of including technical data/characteristics in our models, it has been chosen to perform a likelihood-ratio test. The statistic of this test (*LR test*) and its corresponding *P. value* are reported in Tables 5–7. The line *Nested model* indicates against which model the investigated model is tested. In econometric terms, the nested model is the restricted model and corresponds to the reduced model without any technical data/characteristics.

We turn now to the comments of the results obtained for each sample and sub-sample. We only focus on the signs and significance of the estimated coefficients since the absolute magnitudes of those coefficients are not important. The effects of factors influencing the estimates are then discussed by comparing them with relevant literature, as far as possible.

4.2.1. *Results for the whole sample*

Estimates results for the "*whole*" sample are presented in Table 4, columns (1aAll) and (2aAll). Eq. (9) is estimated using both the White's HCCM (column (1aAll), Table 4) and the WLS (column (2aAll), Table 4) estimators. Contrary to economic primary studies, variances are usually not reported for each estimate in LCA primary studies and have to be retrieved (recall Section 3.1.3). For each observation of the MRA, variances have been directly inserted in the database or calculated depending whether the observations were coming from primary studies performing Monte-Carlo analysis or sensitivity analysis, respectively. As a consequence, the database does not provide a single measure of the variance for each observation. For this reason we prefer to comment coefficient estimates obtained by OLS estimator with a White procedure—*OLS* (*White's HCCM*), as indicated in the last line, Table 4 – rather than *WLS*. However, we present WLS estimates to check for robustness since they yield to similar results. For simplicity's sake, the same choice is applied to the remainder of the paper.

Table 6
Results of MRA for the econometric samples G2-BtL biofuels.

Samples **Model**	**BtL** **1aBtL**	**BtL** **2aBtL**	**BtL** **1bBtL**	**BtL** **2bBtL**	**BtL** **1cBtL**	**BtL** **2cBtL**	**BtL** **1dBtL**	**BtL** **2dBtL**	**BtL** **1eBtL**	**BtL** **2eBtL**
Constant	43.23*** (12.66)	29.68 (20.05)	70.94*** (9.38)	75.11*** (12.85)	39.6*** (6.29)	32.19*** (6.36)	27.43*** (5.59)	29.03*** (5.69)	16.03** (6.33)	24.5*** (8.38)
Technical data										
mat_cultxdluc	−14.41*** (3.08)	−19.53*** (3.27)	−15.41*** (3.11)	−20.33*** (3.23)	−15.16*** (2.93)	−18.1*** (2.84)	−16.24*** (2.47)	−17.02*** (2.6)	−11.64*** (3.98)	−13.28*** (4.46)
cop_elec	−44.56*** (4.14)	−35.68*** (7.7)	−43.99*** (4.19)	−35.94*** (7.71)	−19.4*** (5.09)	–	–	–		
g2_mass_yield			–	–						
g2_mass_yield_sq										
g2_mass_yield_ln	−11.66* (7.04)	−17.41** (8.52)								
btl_pro_autoth										
btl_pro_alng	17.04*** (5.4)	–	17.73*** (5.73)	–	13.85** (5.6)	–				
btl_pro_alelec	–	–	–	–	–	–				
btl_pro_alrenew	–	–	–	–	–	–				
btl_gasrecycl	–	–	–	–	–	–				
Methodological choices										
lca_att (ref)										
lca_cons	25.61*** (8.36)	–	26.23*** (8.6)	–	22.97*** (8.68)	–	18.1** (8.87)	–	–	–
copval_alloc	9.6*** (3.37)	–	9.81*** (3.4)	–	13.73*** (3.05)	12.51*** (4.33)	16*** (2.96)	11.02** (4.65)	11.7*** (2.79)	13.21*** (4.31)
copval_systexp (ref)										
copval_hyb	–	–	–	–	–	–	–	–	–	–
luc_indir	–	–	–	–	–	–	–	–	–	–
uncer_MC	–	–	–	–	–	–	–	–	–	–
uncer_SA	–	–	–	–	–	–	–	–	–	–
uncer_ref (ref)										
impcat_nev	–	–	–	–	–	–	–	–	–	–
impcat_nrc	–	–	–	–	–	–	–	–	–	–
impcat_other	–	–	–	–	–	–	–	–	–	–
impcat_gwponly (ref)										
Typology of the study										
zlab_us	−21.42*** (4.8)	–	−24.38*** (4.71)	−16.37* (9.57)	−22.11*** (4.35)	−16.59*** (5.51)	−16.51*** (4.09)	−14.91*** (5.03)	−11.41** (4.93)	−17.34*** (5.29)
zlab_eu (ref)										
Model information										
N	132	132	132	132	141	141	143	143	143	143
Mean dep. Var.	19.45	21.96	19.45	21.84	18.80	22.29	18.65	22.22	18.65	21.62
Adj. R-squ.	39.48%	26.01%	38.39%	23.56%	35.19%	25.06%	31.39%	25.56%	33.22%	24.68%
Log-Likelihood	−548.53	−568.94	−549.72	−571.09	−589.14	−608.32	−603.08	−618.57	−599.04	−617.29
F-stat. (*P*. value)							16 (0.0000)	15.66 (0.0000)	11.34 (0.0000)	10.69 (0,0000)
Skewness *P*. value)	12.64 (0.2445)		13.7 (0.1869)		13.31 (0.1492)		12.09 (0.0336)		16.91 (0.0501)	
Kurtosis (*P*. value)	2.53 (0.1117)		2.38 (0.1232)		2.22 (0.1364)		1.64 (0.2004)		1.52 (0.2184)	
AIC	1115.07	1155.89	1117.44	1160.18	1194.29	1232.65	1218.17	1249.13	1218.07	1254.57
BIC	1141.01	1181.83	1143.38	1186.12	1217.88	1256.24	1235.95	1266.91	1247.70	1284.20
LR test (*P*. value) Nested model: model (d)	109.1 (0.0000)	99.25 (0.0000)	106.73 (0.0000)	94.96 (0.0000)	27.88 (0.0000)	20.49 (0.0000)			8.09 (0.0882)	2.56 (0.6336)
Procedure	OLS (White's HCCM)	WLS	OLS (White's HCCM)	WLS	OLS (White's HCCM)	WLS	OLS (White's HCCM)	WLS	OLS (White's HCCM)	WLS

Thus, we only comment results presented in column (1aAll), Table 4. 533 observations are included in this regression. As already explained in the previous Section, this regression only aims at testing the influence of (i) the type of biofuels (*gen_3, etha* and *btl* variables) and (ii) the geographical location (*zlab_us*, *zlab_eu* and *zlab_other*) on the *e-s* in order to confirm or deny what have been highlighted with the visual inspections presented in Sections 3.2.1 and 3.2.2. This may explain the rather low level of the adjusted R-squared (about 16%). As judged by the F-stat. *P*. value, the joint significance of results is accepted at the 1% significance level.

As a first comment, the econometric results displayed in Table 4 tend to confirm insights presented in Section 3.2, which were based on a simple visual inspection. *etha* and *btl* variables are indeed statistically significant at the 1% level and their coefficients are negative. According to these parameter estimates, GHG emissions are statistically lower for Ethanol and BtL (G2 biofuels) than for G3 biofuels (*gen_3*) by approximately 41 and 52 g CO_2eq/MJ respectively. These results also confirm that life cycle GHG emission performance is better for BtL than for Ethanol. One cannot effectively merge the *etha* and *btl* variables, as indicated by the Wald Test: we effectively reject the null hypothesis of this test, H_0, because $P.Value < 0.01$ and conclude that the coefficient of *etha* is statistically different from the one of *btl*. Hence the biofuel generation is a key variable to explain the variability of advanced biofuels LCA results.

Regarding the geographical location, *zlab_us* and *zlab_other* variables have a negative impact on GHG emissions – their

Table 7
Results of MRA for the econometric samples G3 biofuels.

Samples Model	G3 1aG3	G3 2aG3	G3 1bG3	G3 2bG3	G3 1cG3	G3 2cG3	G3 1dG3	G3 2dG3	G3 1eG3	G3 2eG3
Constant	318.44*** (90.09)	550.41*** (171.08)	621.72*** (99.61)	916.55*** (146.59)	490.82*** (87.74)	640.23*** (68.43)	450.73*** (88.11)	585.43*** (59.96)	105.38*** (19.91)	237.46*** (25.17)
Technical data										
fame										
hao	134.18*** (35.34)	185.12*** (39.47)	135.18*** (32.44)	181.47*** (34.94)	137.73*** (34.09)	178.64*** (35.28)	134.31*** (34.62)	176.78*** (34.69)		
g3_productivity					–	−5.82*** (1.86)	–	−3.19*** (1.2)		
g3_productivity_sq					–	0.02** (0.01)				
g3_productivity_ln	−65.31*** (20.13)	−124.8*** (45.43)	−64.46*** (20.06)	−127.33*** (44.5)						
g3_oil			−434.74*** (142.4)	−521.06*** (112.68)	−425.28*** (150.9)	−522.41*** (110.74)	−430.6*** (146.08)	−527*** (115.09)		
g3_oil_sq										
g3_oil_ln	−140.32*** (42.3)	−161.66*** (39.17)								
g3_Oppond	−197.13*** (34.6)	−259.9*** (24.28)	−198.94*** (36.92)	−260.33*** (24.35)	−201.93*** (38.32)	−257.79*** (24.13)	−199.6*** (38.46)	−257.1*** (23.73)		
Methodological choices										
lca_att (ref)										
lca_cons	172.72*** (61.08)	250.7*** (70.3)	174.8*** (65.98)	254.66*** (73.36)	196.92** (79.77)	268.68*** (81.63)	187.41** (84.67)	290.31*** (86.21)	–	–
Typology of the study										
zlab_us	−207.56*** (31.01)	−259.02*** (26.35)	−198.73*** (29.55)	−244.44*** (25.73)	−201.53*** (30.19)	−244.11*** (26.09)	−199.27*** (30.8)	−240.76*** (25.61)	−95.93*** (26.17)	−225.96*** (31.21)
zlab_eu (ref)										
zlab_other	–	–	–	–	–	–	–	–	−114.8*** (21.87)	−246.87*** (26.35)
Model information										
N	68	68	68	68	68	68	68	68	69	69
Mean dep. Var.	59.97	68.95	59.97	67.95	59.97	67.85	59.97	66.53	58.84	126.49
Adj. R-squ.	65.23%	80.63%	66.07%	81.32%	66.59%	81.92%	66.06%	81.34%	17.77%	48.39%
Log-Likelihood	−373.01	−376.38	−372.18	−375.14	−371.08	−373.46	−372.19	−375.11	−410.22	−418.31
F-stat. (*P*. value)	11.11 (0,0000)	24.7 (0,0000)	11.67 (0,0000)	25.07 (0,0000)	10.96 (0,0000)	22.62 (0,0000)	13.53 (0,0000)	27.17 (0,0000)	11.17 (0,0000)	31.42 (0,0000)
Skewness (*P*. value)	9.25 (0.2352)		13.93 (0.0524)		14.15 (0.078)		16.85 (0.0184)		32.57 (0,0000)	
Kurtosis (*P*. value)	0.06 (0.8139)		0.32 (0.5694)		0.47 (0.4938)		0.41 (0.521)		6.01 (0.0142)	
AIC	762.02	768.75	760.36	766.29	760.17	764.92	760.37	766.23	828.44	844.62
BIC	779.78	786.51	778.11	784.04	780.14	784.89	778.13	783.98	837.37	853.56
LR test (*P*. value) Nested model: model (e)	74.42 (0,0000)	83.87 (0,0000)	76.08 (0,0000)	86.34 (0)	78.27 (0,0000)	89.71 (0,0000)	76.06 (0,0000)	86.4 (0,0000)		
Procedure	OLS (White's HCCM)	WLS	OLS (White's HCCM)	WLS	OLS (White's HCCM)	WLS	OLS (White's HCCM)	WLS	OLS (White's HCCM)	WLS

coefficients are significant at the 1% level. According to these results, GHG emissions are statistically lower when studies are from NA or from other countries (excluding NA and Europe) compared to those from Europe. Hence, the geographical location appears to have an influence on GHG emission results for advanced biofuels. There is no intuitive reason to explain the geographical influence highlighted by our results. At this step of the analysis, this result could be explained by either a model misspecification or the existence of a publication bias. The former could correspond to missing variables in our database, hence the geographical location could be a shadow variable hiding a *real* determinant. For instance, the geographical location variable could hide a set of technical data specific to one location. Unfortunately, it is not possible to include such variables in the “*whole*” sample model. To test this hypothesis, the “*whole*” sample is thus divided into G3 biofuel sample and G2 biofuel sample in order to assess specific characteristics (including technical data) of each biofuel generation.

4.2.2. Results for the G2 sample

Estimates results for the “*G2*” sample are presented in Table 4, columns (1aG2) to (2bG2). Our comments are based on results presented in column (1aG2). The adjusted R-squared is now approximately 37%.

4.2.2.1. Technical variables. etha variable is statistically significant at the 1% level and impacts positively GHG emissions for G2 biofuels. Thus GHG emissions are higher by about 6 g CO_2eq/MJ for Ethanol than for BtL. The type of fuel conversion technology can thus explain the variability of GHG emission results for G2 biofuels. G2 sample is then split into “*G2-Ethanol*” sample and “*G2-BtL*” samples in order to take into account specificities of each fuel (see Sections 4.2.3 and 4.2.4, respectively).

Regarding the influence of *mat_cult*, this variable was tested first and had a negative effect on GHG emissions for G2 (results reported in columns (1bG2) and (2bG2), Table 4). Most LCA studies do not account for upstream burdens related to residue production

and cultivated feedstock needs more inputs (especially fertilizers and pesticides) to be produced [4] so this result was unexpected. However, it is also well known that perennial energy crops can stock carbon underground [96]. Therefore, our counter-intuitive result could be explained by this fact, but only if direct LUC is accounted for (accounting for above ground and underground carbon sequestration). However we noticed that *luc_dir* variable is not statistically significant. Hence, we decided to combine the *mat_cult* variable with the luc_dir variable (aggregated in *mat_cultxdluc*) in order to confirm this effect (results reported in columns (1aG2) and (2aG2), Table 4). Our meta-model shows that *mat_cultxdluc* variable is statistically significant at the 1% level and impacts negatively GHG emissions for G2 biofuels. It means that GHG emissions for G2 biofuels produced from cultivated feedstock that take into account dLUC are lower than GHG emissions for G2 biofuels from cultivated feedstock that do not take into account dLUC or from waste feedstock. Thus, the type of feedstock combined with the fact that authors take into account dLUC influence GHG emissions for G2 biofuels.

4.2.2.2. Methodological variables. *lca_cons* variable is statistically significant at the 1% level for the "*G2*" sample. Its coefficient is negative so GHG emissions for G2 biofuels are lower with a consequential approach compared to the attributional approach. The type of LCA approach thus influences GHG emission results for G2 biofuels.

copval_alloc and *copval_hyb* variables are statistically significant at the 1% and 5% level, respectively (column (1aG2), Table 4). It confirms the influence of the method for taking into account coproducts on LCA GHG emission results as often mentioned in the literature [64]. The coefficients of both variables are positive which means that GHG emissions are lower for G2 biofuels when using the system boundaries expansion method (*copval_systexp*) compared to allocation and hybrid methods. We observed, however, that most LCA authors recognize the importance of the method applied to account for burdens associated to coproducts. 91% of the studies in our database test alternative methods for allocation performing a sensibility analysis.

luc_indir is statistically significant at the 1% level. It shall be noticed that all studies assessing indirect LUC (*luc_indir*) assess also direct LUC (*luc_dir*), so *luc_indir* is equal to 1 when the study assesses both direct and indirect LUC. Nevertheless *luc_dir* is not statistically significant. We can then conclude that assessing indirect LUC increases GHG emission results for G2 biofuels as *luc_indir* coefficient is positive. Nevertheless, the direct LUC (*luc_dir*) has an influence but it is linked with the type of biomass feedstock used, as mentioned before.

impcat_nev, *impcat_nrc* variables are both statistically significant at the 1% level. The type of other environmental indicators than GHG emissions assessed in the study thus could influence GHG emission results for G2 biofuels. According to our results, GHG emissions are statistically lower when the study assesses the Net Energy Value (*impcat_nev*) and are statistically higher when the study assesses the Non Renewable Energy consumption (*impcat_nrc*). This effect could not have been anticipated. Nevertheless, these variables can be interpreted as a quality indicator for the study: when these energy indicators are consistently assessed, the GHG emission result can be considered to be more robust.

Variables related to the methods for taking into account uncertainties (*uncer_MC* and *uncer_SA*) are statistically significant and impact positively the amount of GHG emissions emitted for G2 biofuels. This effect is unexpected. It means that GHG emissions for G2 biofuels are statistically higher when uncertainties are taken into account—*via* Monte Carlo method (*uncer_MC*) or Sensitivity analysis (*uncer_SA*) than when there is no uncertainties assessment (*uncer_ref*). The assessment of uncertainties by study authors' can also be interpreted as a quality indicator of a study. It can be seen as an effort to establish the accuracy of the results but the tendency of the influence of these parameters in the e-s could not be anticipated nor explained afterward.

4.2.2.3. Typological variables. Lastly, *zlab_us* variable has a negative impact on GHG emissions and is significant at the 1% level. Again, GHG emissions for G2 appear to be statistically lower when the authors are from North America (*zlab_us*) compared to authors from Europe (*zlab_eu*). Hence, the geographical location of the authors also influences GHG emission results for G2 biofuels.

4.2.3. *Results for the Ethanol sample*

Estimates results for the "*G2-Ethanol*" sample are presented in Table 5. Columns (1cEtha) and (2cEtha) correspond to the model without the inclusion of the technical variable representing the mass yield of the pathway (*g2_mass_yield*). Columns (1bEtha) and (2bEtha) test the existence of a linear effect of this variable (*g2_mass_yield*) whereas columns (1aEtha) and (2aEtha) test the existence of a non-linear effect of this variable by taking the logarithm of the *g2_mass_yield* variable (*g2_mass_yield_ln*). The AIC and the BIC both increase from the first columns ((1aEtha) and (2aEtha)) to the last ones (columns (1cEtha) and (2cEtha)). Therefore, the inclusion of a non-linear effect of the mass yield of the pathway appears more relevant to explain GHG emission variations. Thus, we choose to comment results presented in column (1aEtha).

4.2.3.1. Technical variables. *mat_cultxdluc* variable is significant at the 5% level and has the same effect on GHG emissions for Ethanol as for G2 biofuels (see Section 4.2.2).

The mass yield of the pathway *g2_mass_yield_ln* impacts negatively GHG emissions for G2 Ethanol, which is an intuitive effect: the better the mass yield is, the less GHG are emitted all along the biofuel life cycle, ceteris paribus. It should be noticed that *g2_mass_yield_ln* traduces a non-linear effect of this variable.

We should also mention that variables related to other technical data, such as the type of biomass pretreatment, are not statistically significant for Ethanol. Indeed, 83% of observations are related to Ethanol produced using dilute sulfuric acid pretreatment and most of these observations use technical data from the same study (NREL) [97]. Hence pretreatment process variables for Ethanol are not really discriminatory, and this could explain why those variables are not statistically significant.

4.2.3.2. Methodological variables. Among significant variables found for G2 biofuel sample, *lca_cons, luc_indir, impcat_nev, impcat_nrc, uncer_MC* and *uncer_SA* are also significant for the Ethanol sample and have the same impact as described for the G2 sample. So the type of LCA approach, the fact 'to assess indirect LUC', the type of other environmental indicators, the method for taking into account uncertainties influence GHG emission results for G2 Ethanol.

It can be noticed that *copval_alloc* and *copval_sys* variables are no longer statistically significant. This result is surprising regarding a previous lignocellulosic Ethanol LCA studies review [14] which concludes that the treatment of coproducts has a strong influence in the LCA results.

4.2.3.3. Typological variables. *zlab_us* variable has a negative impact on GHG emissions and is significant at the 5% level. It means that GHG emissions for Ethanol are statistically lower when the authors are from North America (*zlab_us*) compared to authors from Europe (*zlab_eu*). Hence, the geographical location of the authors also influences GHG emission results for Ethanol.

4.2.4. Results for the BtL sample

Estimates results for the "*G2-BtL*" sample are presented in Table 6. Columns (1eBtL) and (2eBtL) correspond to the reduced model obtained for the "*G2*" sample. Columns (1dBtL) and (2dBtL) correspond to the *new* reduced model without technical variables. Columns (1aBtL) to (2cBtL) correspond to the reduced model with technical variables. Columns (1aBtL) and (2aBtl) are the only ones to test a non-linear effect of the mass yield of the pathway. The AIC and the BIC both increase from the first columns ((1aBtL) and (2aBtL)) to the last ones (columns (1eBtL) and (2eBtL)). Thus, we choose to comment results presented in column (1aBtL).

4.2.4.1. Technical variables. *mat_cultxdluc* variable is significant at the 1% level and has the same effect on GHG emissions for BtL as for G2 biofuels (see Section 4.2.2)

Variables related to the type of fuel conversion process (*btl_pro_alng* and *btl_pro_alelec*) are statistically significant. Using natural gas as a source of heat for an allothermic BtL unit leads to higher GHG emissions than producing BtL from an autothermic plant (biomass provides all process energy needs). Conversely using grid electricity as a utility for an allothermic BtL unit leads to lower GHG emissions than producing BtL from an autothermic plant. The source of electricity used could explain these results. Indeed, among the observations using grid electricity as a utility for an allothermic BtL unit, 57% of these observations use electricity provided by wind power plants [62]. The other studies do not precise the source of electricity used.

The mass yield of the pathway *g2_mass_yield_ln* impacts negatively GHG emissions for BtL, which is an expected effect: the better the mass yield is, the less GHG emissions are emitted all along the pathway for a G2 biofuel. It should also be noticed that *g2_mass_yield_ln* traduces a non-linear effect of this variable.

Variables related to other technical data, such as the type of biomass pretreatment or the inclusion of Carbon Capture and Storage (CCS) in the process, are not statistically significant for BtL. Indeed, 90% of the observations in the econometric sample are related to BtL produced without biomass pretreatment (see Table II.4 in the Supplementary Data). Hence pretreatment process variables for BtL are not really discriminatory, and this may be the reason why those variables are not statistically significant. Moreover, the variable *btl_ccs* is equal to zero in the econometric sample (see Table II.4 in the Supplementary Data), therefore this variable could not have been tested. In fact, the variable in question appears in only three observations and all of them are considered outliers (see Table I.8 in the Supplementary Data).

4.2.4.2. Methodological variables. Among significant variables found for the "*G2*" sample, only *copval_alloc* and *lca_cons* are significant for the "*G2-BtL*" sample. The method for taking into account coproducts (*copval_alloc*) has the same impact as described for "*G2*" sample (*copval_hyb* for BtL is equal to zero). However the influence of the type of LCA approach is not the same for G2 biofuel and for BtL: GHG emissions are higher with a consequential approach (*lca_cons*) compared to an attributional approach (*lca_att*). So the type of LCA approach and the method for taking into account coproducts influence GHG emission results for BtL.

Furthermore, the type of coproduct influence GHG emission results for BtL since the *cop_elec* variable is statistically significant at the 1% level. Therefore, the coproduction of electricity in a BtL production plant decreases life-cycle GHG emissions compared to other coproducts, ceteris paribus.

4.2.4.3. Typological variables. *zlab_us* variable has a negative impact on GHG emissions and is significant at the 1% level. It means that GHG emissions for BtL are statistically lower when the authors are from North America (*zlab_us*) compared to authors from Europe (*zlab_eu*). Hence, the geographical location of the authors also influences GHG emission results for BtL.

4.2.5. Results for the G3 sample

Estimate results for the "*G3*" sample are presented in Table 7. We begin by commenting the impact of *g3_productivity* and *g3_oil* as the influence of these two continuous technical variables will determine the final specification of the model for the "*G3*" sample.

4.2.5.1. Technical variables. First, a lin-lin model is specified in order to test the linear effects of both *g3_productivity* and *g3_oil* on the *e-s*. Table 7, column (1dG3) shows the reduced form of this specification. It can be noticed that the *g3_productivity* variable is not statistically significant. This result is non-intuitive as most of the literature mentions that algae productivity can explain the variability of GHG emission results. The non-significance of this variable may be explained by the existence of a non-linear effect instead of a linear one. To test this hypothesis, two models are specified. In the first one (Table 7, column (1cG3)), the non-linear effect is modeled as a second-degree polynomial by introducing the variable *g3_productivity* and its squared value (*g3_productivity_sq*). In the second one (Table 7, column (1bG3)), the linear effect is modeled as a logarithmic function by introducing *g3_productivity_ln* instead of *g3_productivity*. In Table 7, column (1cG3), neither *g3_productivity* nor *g3_productivity_sq* are statistically significant at the 10% level. On the contrary, *g3_productivity_ln* is statistically significant at the 1% level (Table 7, column (1bG3)). As a conclusion, the variable *g3_productivity* does have an impact on GHG emission results for G3 biofuels but its effect is non-linear, which can be captured by a logarithmic function, not a second-degree polynomial. Regarding *g3_oil*, results presented in Table 7, column (1bG3), indicate a negative linear influence of this variable. Finally only *g3_productivity_ln* and *g3_oil_ln* variables are statistically significant at the 1% level and their coefficients are both negative (Table 7, column (1aG3)). Thus, we choose to comment the results presented in column (1aG3).

Algae productivity value and the oil content – as proxies of the *g3_productivity* and *g3_oil* variables, respectively—influence GHG emission results for G3 biofuels. They have a negative impact on GHG emissions so the higher the algae productivity or the algae oil content is, the lower the GHG emissions are. In addition, these non-linear effects indicate that those parameters are more sensitive for low productivity or low oil content than for high ones.

The variable *hao* is statistically significant at the 1% level. According to its coefficient estimate, GHG emissions for HAO from algae are higher than GHG emissions from FAME from algae by about 134 g CO_2eq/MJ ceteris paribus. It indicates that the type of fuel conversion technology can explain the variability of GHG emission results for G3 biofuels. This result is difficult to be interpreted, especially due to the extent of its coefficient. In fact, the literature shows that upstream fossil energy consumption (including all inputs, notably methanol and hydrogen production) is similar in FAME and HAO processes [98],[14]. Algal oil consumption for both processes is also quite similar.

The coefficient of *g3_Oppond* is negative and significant at the 1% level. GHG emissions for G3 biofuels are thus statistically lower when microalgae are grown in an open-pound than in a

[14] In the EcoInvent database [119], the cumulative fossil energy demand for the production of 1 kg of hydrogen from cracking natural gas is 70, 9 MJ. The same indicator for 1 kg of methanol also produced from natural gas is 36.9 MJ. FAME contains around 10% of methanol and HAO around 4% of hydrogen (mass). The selected processes for this example are the most commonly used for these products.

photobioreactor. Hence the type of technology used for microalgae cultivation influences GHG emission results for G3 biofuels.

The type of technology used for microalgae cultivation, the algae productivity and the oil content of algae are often identified as key parameters in G3 biofuel LCA studies. So the fact that those variables are statistically significant confirms previous conclusions found in the literature. Jorquera et al. [99], in a microalgae LCA study (not included in this review because conversion into biofuel is not included), shows that culture in photobioreactors is more energy intensive than in open ponds. One of the conclusions of previous literature reviews on microalgae biofuel technologies [100,101] is that microalgae strains presenting high biomass productivity are better for CO_2 emission mitigation.

4.2.5.2. Methodological variables. Concerning methodological variables, only *lca_cons* variable is statistically significant at the 1% level for the G3 sample. Its positive coefficient indicates that GHG emissions for G3 biofuels are statistically higher when the study uses a consequential approach for LCA compared to the attributional approach. Hence the type of LCA approach influences GHG emission results for G3 biofuels. However, note that consequential LCA approach is only used by one study (that represents 9% of the observations for the econometric sample). Consequently the influence of the type of LCA approach for G3 biofuels should be interpreted with caution.

Liu et al. [18], in an LCA harmonization exercise, show that different authors accounted for different microalgae coproducts and that it plays an important role in the final life cycle GHG emissions of the biofuel. However, in our meta-regression, variables related to the coproducts did not show themselves to be statistically significant. Still, the fact that *lca_cons* is statistically significant warns us about the importance of the definition of system boundaries and coproduct accounting methodology.

4.2.5.3. Typological variables. Regarding typological variables, the coefficient of *zlab_us* variable is significant at the 1% level and its sign is negative whereas *zlab_other* is not statistically significant. Thus, the previous result regarding the influence of geographical location is partly retrieved: GHG emissions of G3 biofuels are statistically lower when studies are from NA compared to ones from Europe. The non-significance of *zlab_other* indicates that there are no systematic differences between results drawn from European studies and other countries.

4.2.6. Discussion on MRA results

The MRA results presented in Sections 4.2.1–4.2.5 indicate that life-cycle GHG emissions of G3 biofuels are statistically higher than those of Ethanol which, in turn, are higher than those of BtL. It confirms the influence of the type of biofuel to explain the variability of advanced biofuel GHG emissions as deduced from the descriptive statistics in 3.2. Additionnaly, the results from North-American studies are statistically higher than the results from European studies. There is no intuitive reason to explain this geographical influence highlighted by our results. It could be explained by either a model misspecification or the existence of a publication bias.

The methodological choices that can influence the LCA results were also identified. Some of those variables are often mentionned in the literature such as the type of LCA approach (A-LCA vs. C-LCA) [102], the method to account for coproducts [20–22] and the inclusion of iLUC [103]. However, the MRA reveals that some non-intuitive variables also influence the results, such as the type of uncertainty analysis conducted in the study or the number of environnmental indicators assessed. A deeper work should be conducted to understand the reasons why such variables influence the results, especially to check if there is no *shadow* variable that would explain this influence.

Moreover, results concerning the technical variables that have an influence on GHG emission estimates were drawn from the MRA. The mass yield has a negative and non-linear effect for both Ethanol and BtL. In the analyzed sample, the type of process has a statistically significant effect only for BtL. The type of biomass fed into the conversion unit is also an influencing variable for G2 biofuels. These variables are often mentioned in the literature as key variabes influencing GHG emission estimates [e.g. 75, 82]. With respect to G3 biofuels, the algae productivity and its oil content have systematically a negative and non-linear effect on the LCA results. Also, the type of technology used for microalgae cultivation influences GHG emissions estimates. G3 biofuel LCA studies also highlight these variables to explain the variability of GHG emissions [99–101]. Nevertheless, the reason why some identified variables influence the results remains unclear (e.g. the type of G3 biofuel conversion—FAME or HAO).

Finally, conclusions can also be drawn from important variables mentionned in the literature that have not been identified by the MRA as variables influencing the final LCA result— for example, the type of biomass pretreatment in the Ethanol conversion process and the use of CCS in the BtL conversion process. The former is probably not statistically significant because most of the Ethanol technical data used in the different studies are derived from one single study [97]. The latter is a variable expected to have a negative impact in the GHG emission results but that could not be tested because all observations with the use of CCS were cut out from the original sample (they were all negative abelow the 5^{th} percentile).

4.3. Harmonization

The MRA results presented in Section 4.2 are now used to address the harmonization issue in the field of advanced biofuels GHG emissions thanks to the technique of *benefits transfer using meta-regression models*. As already demonstrated in the previous section, the *meta-regression* framework allows the production of an estimation of the mean *e-s* weighted by the systematic influence of its main drivers. Once estimated, the meta-function can be used to deduce original values of the *e-s* by specifying new values for the main drivers identified corresponding to relevant case studies. This technique of *benefits transfer using meta-regression models*, as it is named in the MA literature, may be a particularly well adapted methodology to deal with the so-called harmonization issue specific to the LCA literature.

This section aims at providing an illustration of the potential for MRA to perform harmonization in the field of LCA through an application to advanced biofuels GHG emissions. To do so, predicted values of the *e-s* are computed using the meta-functions estimated in Section 4.2.

The predicted values can be calculated using a combination of variables that already exists in the meta-database: this type of prediction is called “in sample”. In sample prediction enables the comparison of collected values (estimations of the *e-s*) and predicted values in order to check the accuracy of the meta-function in predicting the *e-s*.

Furthermore, predicted values can be extrapolated for a combination of relevant variables that do not necessarily exist in the meta-database, hence the prediction is called “out of sample”. Out of sample prediction could provide values for the *e-s* for case studies not assessed in the literature. In addition, out of sample prediction applied to quantitative variables can help to test how sensible the *e-s* is to these variables.

First, in sample predictions are presented and analyzed. Second, out of sample predictions are conducted assessing in particular the sensitivity of quantitative variables (algae productivity and oil content for G3 biofuels, mass yield for BtL and Ethanol).

Table 8
Characteristics of collected and predicted values of the *e-s* in g CO_2eq/MJ (predicted values calculated from (1a) meta-models).

Samples	Whole	G3	G2	BtL	Ethanol
Collected values					
Number of values	533	69	464	143	321
Mean	28.64	58.84	24.15	18.65	26.61
Min	−85.00	−85.00	−24.00	−24.00	−23.65
Max	332.20	332.20	85.80	85.68	85.80
NA values higher than −60% GHG emission threshold	7%	14%	5%	1%	7%
EU values higher than −60% GHG emission threshold	25%	30%	24%	17%	27%
Predicted values					
Number of values	533	68	464	132	209
Mean [confidence interval]	28.64 [25.19;32.09]	59.97 [43.29;76.65]	24.15 [22.56;25.74]	19.45 [16.67;22.23]	19.7 [17.37;22.03]
Min	−9.42	−109.25	−15.82	−8.04	−20.86
Max	76.27	230.82	47.91	56.31	47.49
Underestimated values	44%	46%	44%	47%	47%
Overestimated values	56%	54%	56%	53%	53%
Collected values included in the predicted value CI	12%	51%	22%	18%	37%
NA values higher than −60% GHG emission threshold	5%	9%	2%	1%	2%
EU values higher than −60% GHG emission threshold	7%	28%	22%	8%	1%

4.3.1. Prediction in sample

Table 8 presents some characteristics of predicted values compared to collected values (estimations of the *e-s* in the meta-database) for each sample. The meta-models used to calculate these *in sample predictions* are those estimated in columns (1aAll), (1aG2), (1aEtha), (1aBtL) and (1aG3) for the “*whole*” sample, the “*G2*” sample, the “*G2-Ethanol*” sample, the “*G2-BtL*” sample and the “*G3*” sample respectively (see Tables 4–7).

First, we observe that the mean values for predicted values are slightly different from those of collected values. Nevertheless the ranking between G2 and G3 biofuels, BtL and Ethanol in terms of contribution to the climate change (i.e. amount of GHG emissions emitted all along their life cycle) is still the same as depicted in the econometric analysis. Second, the range of variation is narrower for predicted values than for collected values, except for the G3 sample. Furthermore, these meta-models tend to overestimate predicted values compared to their corresponding collected values (53–56% of predicted values are overestimated depending on the samples) as depicted in Fig. 6.

4.3.2. Prediction out of sample

Out of sample prediction enables the building of values of the *e-s* for combinations of variables that do not necessarily exist in the meta-database. Those values are calculated from the meta-function obtained by the meta-regression method. This harmonization method allows us to obtain mean values of the *e-s* and associated confidence intervals (CI) for each combination of statistically significant variables of a meta-model. For instance, using the meta-model for the “*whole*” sample presented in column (1aAll), Table 4, predicted values of the *e-s* can be calculated for G3 biofuel, BtL and Ethanol in Europe and North America. Table 9 and 10 illustrate the procedure. Table 9 reports coefficient estimates of the model (1aAll) (as presented in column (1aAll), Table 4) and the different values of the variable of this reduced model which have to be imputed to compute the predicted values of the *e-s* for G3 biofuel, BtL and Ethanol in Europe and North America. Table 10 shows the link between these imputed values and the corresponding predicted values of the *e-s* whereas Fig. 7 offers an alternative view of Table 10 results.

As depicted in the Fig. 7, predicted values of GHG emissions for advanced biofuels in Europe are always higher than those in North America. In addition, GHG emissions are lower for BtL than for Ethanol, and G3 biofuels always emit more GHG emissions than G2 biofuels. Those results are in line with the statistical description conducted in Section 3.2. Furthermore, the predicted value CIs are wider for G3 biofuels than for G2 biofuels, meaning that the model better estimates G2 biofuels GHG emissions than those of G3 biofuels. It should be noted that predicted values of GHG emissions for advanced biofuels are always lower than GHG emissions for the reference fossil fuel even when considering CI, except for G3 biofuels in Europe.

The same type of analysis could be conducted for each meta-model. Out of sample prediction could also be used to test the sensitivity of results for quantitative variables. A range of values for quantitative variables could be tested by calculating mean predicted values for the *e-s* and the associated CI, ceteris paribus.

For instance, the influence of oil content and algae productivity is tested for G3 biofuels (Figs. 8 and 9), by testing the range of values found in the meta-database. Results show that both variables have a non-linear effect on LCA GHG emissions, ceteris paribus. Furthermore, variations for high values of the algae productivity have less effect on the *e-s* than variations for low values. Moreover, CIs are smaller for oil content and algae productivity values around mean values than for extreme values.

The same type of sensitivity analysis is conducted to test the influence of the mass yield of the BtL and Ethanol conversion processes on GHG emission results. As depicted in Figs. 10 and 11, the mass yield value has a non-linear effect on LCA GHG emissions of G2 biofuels, ceteris paribus. Variations for high values have less effect on the *e-s* than variations for low values. In addition, CIs are smaller for mass yield values around mean values than for extreme values, as previously described in the G3 sample.

5. Concluding remarks and discussion

This article aims at synthesizing the literature of LCA studies that have estimated GHG emissions of advanced biofuels. Our literature review showed a high variation among the results (Fig. 1). Thus, one can wonder (i) if there is a consensus about GHG emission benefits from advanced biofuels and (ii) why there is so much variation among results. To do so, we have chosen to apply a specific MA methodology (the “*meta-regression anlysis*”, MRA) rather than a more classical narrative literature review approach. It provides a multivariate statistical analysis of previous estimated results to synthesize the available information. This assessment brings an extensive overview and contributes for a better understanding of the main factors inducing GHG emission variations. By using this original quantitative research framework, this article attempts to take the analysis of advanced biofuel GHG emissions one step further by complementing the qualitative

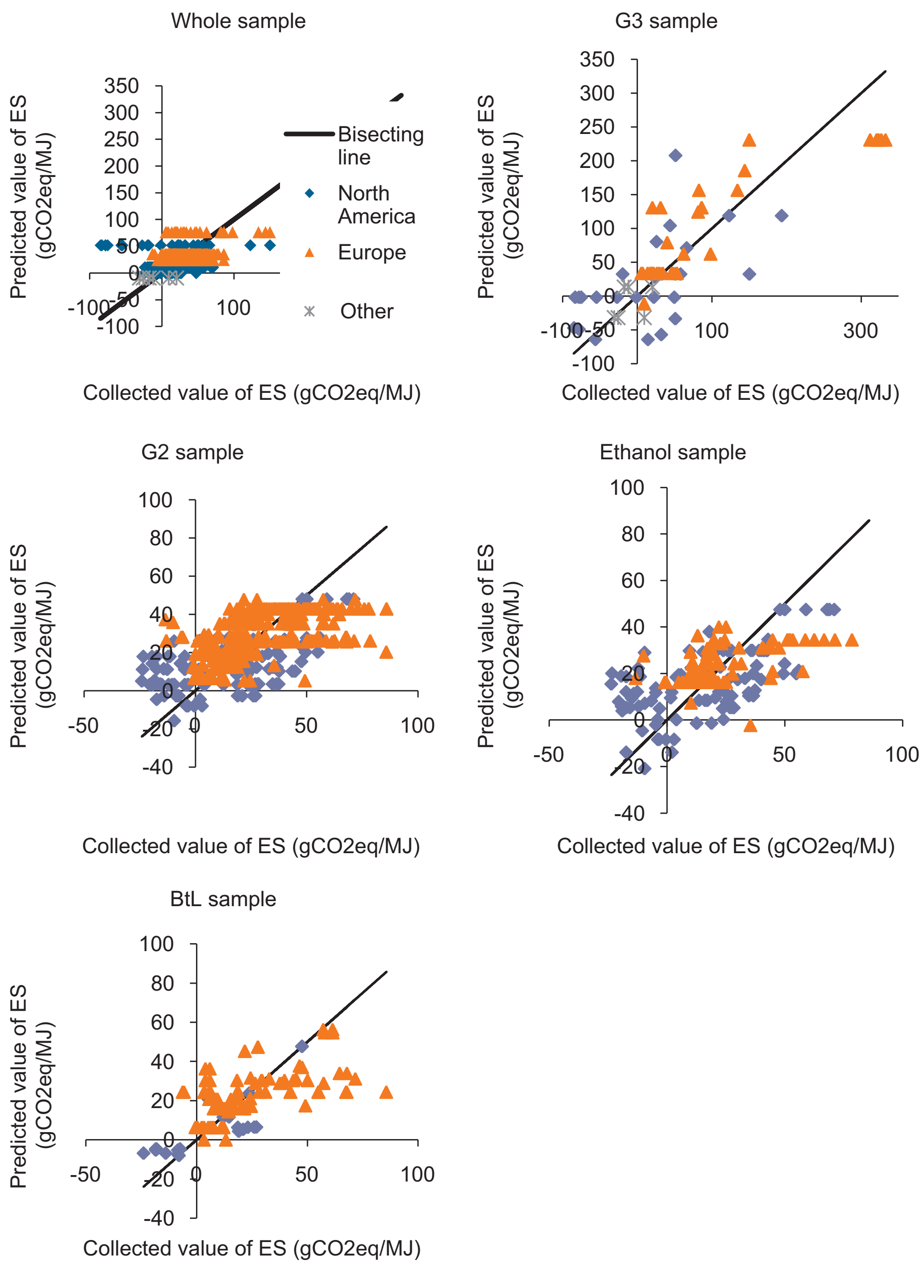


Fig. 6. Predicted and collected values of the *e-s* for meta-model (1a) distinguished by their geographical location.

Table 9
Benefits transfer for the "Whole" sample (1aAll meta-model).

Samples Model: parameter estimate	Whole 1aAll	Imputed values					
Constant	76.27*** (13.64)	1	1	1	1	1	1
Technical data							
gen_3 (ref for Whole)		0	0	0	0	0	0
etha	−41.39*** (13.14)	0	1	0	0	1	0
btl (ref for G2)	−52.12*** (13.36)	0	0	1	0	0	1
Typology of the study							
zlab_us	−24.6*** (3.97)	0	0	0	1	1	1
zlab_eu (ref)		0	0	0	0	0	0
zlab_other	−85.69*** (15.6)	0	0	0	0	0	0
Transfer values		**76,27 (13,64)**	**34,88 (1,75)**	**24,15 (1,88)**	**51,67 (12,35)**	**10,29 (2,98)**	**−0,44 (3,50)**

Table 10
Harmonized *e-s* (g CO_2eq/MJ) for the "Whole" sample (1aAll meta-model)).

Harmonized *e-s*			**95% Confidence interval**	
			Min	**Max**
Europe	**G3**	76.27	49.54	103.00
	G2 Ethanol	34.88	31.45	38.31
	G2 BtL	24.15	20.46	27.84
North America	**G3**	51.67	27.47	75.88
	G2 Ethanol	10.29	4.45	16.13
	G2 BtL	−0.44	−7.31	6.42
In sample predicted value	**Mean**	28.64	25.19	32.09

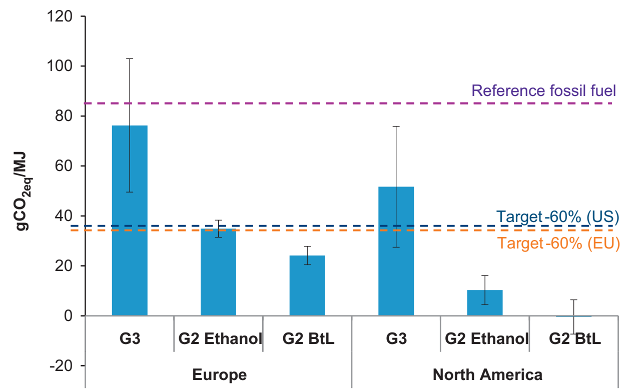


Fig. 7. Predicted values of the effect size for the whole sample calculated from meta-model 1aAll.

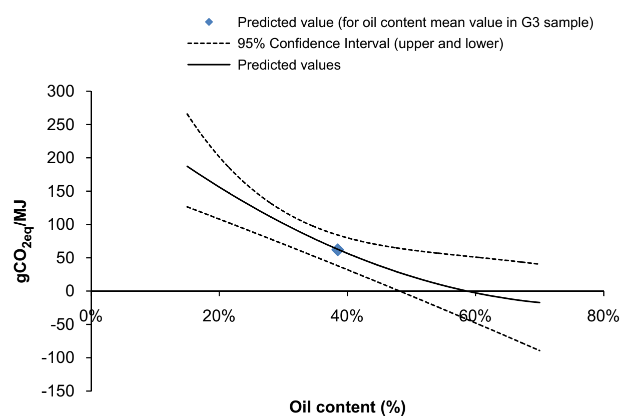


Fig. 8. Influence of oil content on predicted values of the *e-s* for G3 sample (1aG3 meta-model).

surveys which have already been published [9–14]. We investigate through an application the potential for MRA to synthesize LCA literature by highlighting the main determinants of result variability in order to perform harmonization.

Our primary purpose was to identify and quantify which factors among (i) technical data/characteristics, (ii) author's methodological choices and (iii) typology of the study under consideration have an impact on variations of the GHG emission estimates. Our results indicate a hierarchy between G3 and G2 biofuels: GHG emissions of G3 biofuels are statistically higher than those of Ethanol which, in turn, are higher than those of BtL. Moreover, whatever the type of advanced biofuel considered, North-American estimates are statistically higher than European estimates. Regarding author methodological choices, we have shown that some variables can influence the LCA results, such as the type of LCA approach (A-LCA vs. C-LCA), the method to account for coproducts and the fact of taking into account iLUC. Some technical variables

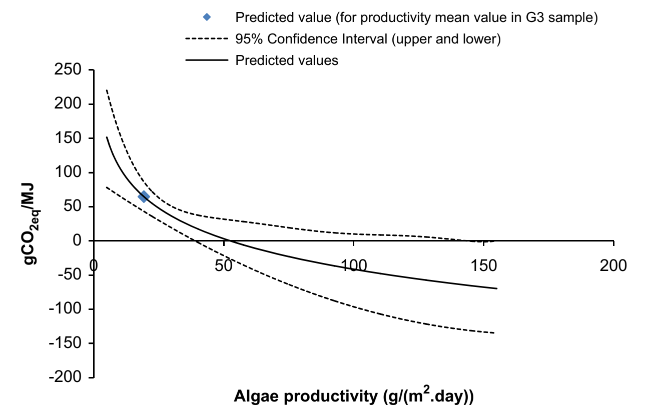


Fig. 9. Influence of algae productivity on predicted values of the *e-s* for G3 sample (1aG3 meta-model).

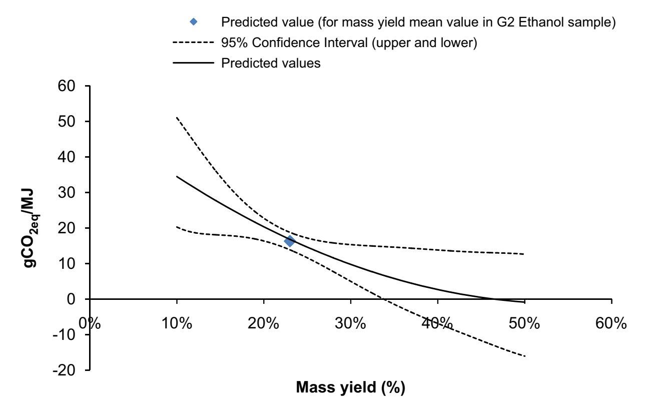


Fig. 10. Influence of the mass yield on predicted values of the *e-s* for Ethanol sample (1aEtha meta-model).

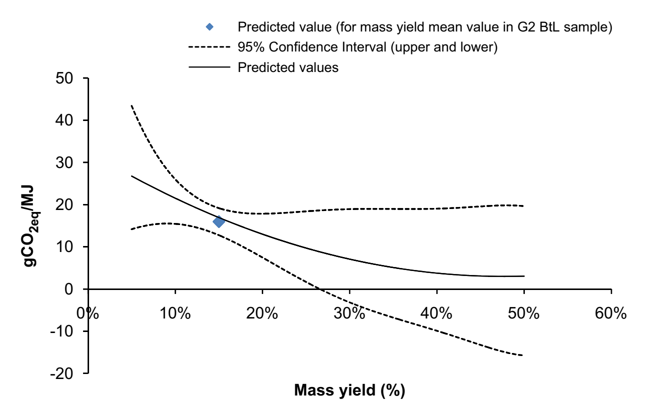


Fig. 11. Influence of the mass yield on predicted values of the *e-s* for BtL sample (1aBtL meta-model).

appear to have an influence on GHG emission estimates. Concerning G2 biofuels, the mass yield has a negative and non-linear effect for both Ethanol and BtL whereas the type of process has a statistically significant effect only for BtL. For G3 biofuels, the algae productivity and its oil content have systematically a negative and non-linear effect. Conclusions can be drawn also for some variables that have not been identified as variables influencing the final LCA result. The the type of biomass pretreatment in the Ethanol conversion process is probably not statistically significant because most of the Ethanol studies in this literature review use data from one single study [97]. The use of CCS in the BtL

conversion process is a variable expected to have a negative impact in the GHG emission results but it could not be tested. All observations with CCS technology fell in the outliers category.

The secondary purpose of this study was to address the harmonization issue in the field of advanced biofuel GHG emissions by using the technique of *benefits transfer using meta-regression models*. Our results may be summarized as follows. For each type of biofuel, a mean value of life cycle GHG emissions (expressed in g CO_2eq/MJ of biofuel) weighted by the influence of its main drivers and its corresponding Confidence Interval is provided (Fig. 5): about 60.0 (ranging from 43.3 to 76.7) for G3 biofuels; 19.7 (ranging from 17.4 to 22.0) for Ethanol; and 19.5 (ranging from 16.7 to 22.2) for BtL. Lastly, these values appear systematically higher for North-American estimates compared to those from Europe, ceteris paribus (Fig. 7). Note that this range of values is lower than the fossil reference (about 83.8 in g CO_2eq/MJ). However, only Ethanol and BtL do comply with the GHG emission reduction thresholds defined in both the US and EU directives.

Some results highlighted in this MRA have revealed some new information not previously assessed in this literature such as the existence of some non-linear effects regarding technical variables. Moreover, MRA provide (i) a measure of the mean *e-s* and (ii) a measure of the precision of this mean value estimate as provided by the corresponding Confidence Intervals. Compared to the only MRA applied to LCA [38], we have gone further by proposing a method to predict LCA results using a meta-model. This can be seen as a statistical harmonization method alternative to the one applied currently in LCA MA using quantitative adjustments as conducted in [18,31–37] for instance.

The common goal of these different MA methodologies is to better understand the main determinants of LCA results in order to propose one mean estimate, also called harmonization in the literature as defined by Heath and Mann [16]. Quantitative adjustment MA [18,31–37] are able to reduce variability in calculated outcomes representing a useful starting point for more precise estimates of LCA results. However, this does not mean that this "harmonization" procedure produces more accurate results since the "more consistent methods and assumptions" applied are subjective. Different authors can consider different methods and assumptions to be more consistent. Conversely, our meta-database is only based on material directly drawn from the literature in order to reduce this kind of subjectivity. The meta-model is obtained from a meta-regression, therefore, it contains the parameters that were statistically proven to influence LCA results in a given sample. Our results show that, with this approach, we can provide more than a mean value and an interquartile range for the e-s. We can calculate a real confidence interval for our predictions.

Furthermore, as highlighted in [39] from 1976 in the biomedical field, "[MRA] *connotes a rigorous alternative to the casual, narrative discussions of research studies which typify our attempts to make sense of the rapidly expanding research literature"*. From our point of view, significant progress can be made in the literature review of LCA studies by applying this methodology and we would recommend that the LCA community should work more closely with the Econometrics community so that more MRA could be conducted.

However, there are many limitations typically associated with MA. In the construction of the database, for example, there is always some exogenous information that has to be provided. Even if we avoid it as much as possible, in some cases it is necessary. This happened especially in the calculation of the e-s where the data required for the conversion of units (LHV, density, motor performance, etc.) was not always provided by the study in question.

Moreover, there is a compromise that has to be made between the number of studies that pass the screening process and the number of independent variables that are used in the description of an observation. In a MA database, all of the observations in a given sample have to be described with the same amount of independent variables. Theoretically, all the parameters that potentially influence the *e-s* have to be included. However, in LCA, the results are affected by hundreds of inputs and methodological choices, making it impossible to fully explain all the results of a big number of observations given the heterogeneity in LCA reporting. It was our judgment and experience in conducting LCA studies, but also previous narrative surveys, that determined which explanatory variables should be included in the database.

Finally, there may be some limitations regarding the statistical population of the MA sample. Heath and Mann [16] highlight the fact that a MA cannot make up for a lack of studies on a certain technology or methodological issue. In our case, for example, there are only 3 observations for BtL including CCS in its production pathway and these were coincidently discarded from the meta-regression sample as outliers. Therefore, no conclusions could be drawn from this technological parameter. Another example is the limited number of consequential LCAs, also limiting the conclusions we can reach concerning this methodological choice.

On our view, MA appears thus more as a complementary methodology than an alternative one to more classical narrative surveys.

Acknowledgments

The authors thank the DDEEP-seminar's participants of EconomiX-CNRS, University of Paris Ouest for their helpful comments as well as to the two anonymous reviewers. The authors also are grateful to Anthony Hua for his valuable help to build the database. The usual disclaimers apply.

Appendix A. Technical description of advanced biofuels

Fig. A.1 represents the main steps involved in the production of second and third generation biofuels (G2 and G3 biofuels respectively) discussed in this paper and the following text contains a brief description of their production processes.

Second generation Ethanol is obtained from the biochemical conversion of annual crop residues (e.g. corn stover) and perennial

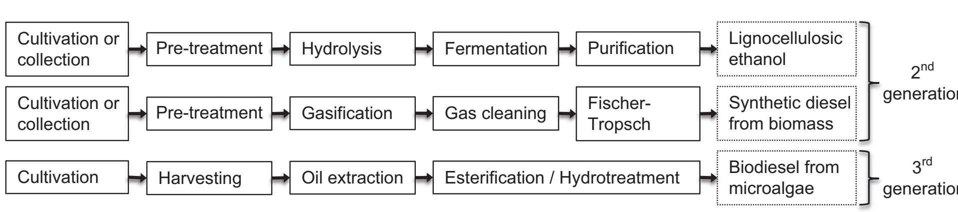


Fig. A.I. Main steps in the production of advanced biofuels.

crops (e.g. miscanthus). A pretreatment of the biomass is necessary to separate the cellulose from hemicellulose and lignin. Once the cellulose is accessible, enzymes are used to hydrolyze these molecules, transforming them into sugars that can be fermented. The product of fermentation needs to be distilled and dehydrated in order to obtain pure Ethanol [97,104].

Synthetic diesel from biomass is also known as Biomass to Liquids (BtL) or biomass FT-diesel. It is produced by the thermochemical conversion of forest residues, herbaceous energy crops (e.g. switchgrass) and woody biomass (e.g. poplar). A pretreatment of the biomass is necessary so that it can be loaded into the gasifier. In the gasifier, the biomass suffers a thermal treatment (partial oxidation) into what is known as "syngas", composed mainly of H_2 and CO. Impurities are removed from the "syngas" during a gas cleaning step, due to the high sensibility of the Fischer–Tropsch (FT) reaction catalyst. The synthetic diesel is obtained after the upgrading (hydrocracking) of the products from the FT unit [105,106].

Biodiesel can be produced from conventional transesterification of oil extracted from microalgae that have a higher theoretical productivity per hectare than conventional vegetable oil crops (e.g. soybeans, palm). Microalgae can be cultivated in open ponds or photobioreactors (PBR) and the technologies for harvesting, drying and extracting oil still require considerable research effort. Various pathways are studied in order to reduce costs and energy consumption in the production process. The use of power plant flue gas as a CO_2 source for growing algae or wastewater as a source of nutrients are potential options for this biodiesel pathway [100,107].

Studies about hydrotreated algal oil (HAO) from the hydrogenation of microalgae oil were also included in this literature review. It has different characteristics than biodiesel but the most important life cycle steps involving microalgae growth, harvesting and oil extraction are the same. HAO, as well as BtL, are being studied as renewable alternatives not just for road transportation but also for the aviation industry.

Appendix B. Complements on the MRA theory

This appendix aims at the explanation of the treatment of heteroskedasticity in MRA.

Heteroskedasticity is a well-known problem in MRA literature. Recall that the basic linear regression model assumes homoskedasticity, i.e. equal variances of ε_i:$E(\varepsilon\varepsilon') = \sigma_\varepsilon^2 I$. This assumption assumes that the variance of the error terms is the same for all observations. It implies that the variance-covariance matrix of the vector of parameters estimates, $\hat{\beta}_{(K,1)}$, is equal to $\hat{\sigma}_\varepsilon^2 (X'X)^{-1}$. More particularly, it is thus assumed: $\sigma_{\varepsilon,i}^2 = \sigma_\varepsilon^2, \forall i = 1, \ldots, I$. When applied to the MRA framework, the homoskedasticity assumption of the disturbances may not be held.

By nature, primary studies results are not estimated with the same precision. In econometric terms, it means that each estimate has a different standard error, that is: $\sigma_{\varepsilon,i} \neq \sigma_{\varepsilon,j}, \forall i \neq j$. As a consequence, the variance of ε in Eq. (1) varies across its observations and *e-s* estimates, y_i, may not be considered as having homogeneous variances. Indeed, "*e-s*" estimates are drawn from different primary studies. These studies use different (i) technical data/characteristics, (ii) author's methodological choices and (iii) do not have the same typology. These reasons, among others, may explain why each *e-s* estimates are estimated with varying degrees of precision.

In presence of heteroskedasticity, the Ordinary Least Square (OLS) estimates, $\hat{\beta}_{(K,1)}$, remain unbiased and consistent. Nevertheless, heteroskedasticity often leads to wider parameter estimate confidence intervals, which may cause insignificant relationships between independent and dependent variables if not accounted for[15]. Therefore, heteroskedasticity is potentially a serious problem and has to be explicitly treated in MRA. Various solutions have been used in the MRA literature to correct for heteroskedasticity[16]. Two majors approaches have been employed in the literature:

Methods of estimation using Heteroskedastic Consistent Covariance Matrix

One of the most common approaches is to use heteroskedastic consistent estimators such as White's or Huber-White's Heteroskedastic Consistent Covariance Matrix (HCCM). The Newey-West estimator has also been used in some MA. The latest has been designed for stationary time-series data and, as a consequence, Nelson and Kennedy [92] do not recommend to employ this estimator in a MRA framework. The use of White and/or Huber-White standard errors theoretically corrects for heteroskedasticity.

Nevertheless, non-homogeneous variances may remain in practice, more particularly when MRA are applied to small sample sizes. The white and Huber-White estimators are generally used because the source of heteroskedasticity is not exactly known. It is not the case in the context of MRA in which the source of heteroskedasticity is clearly identified. Indeed, it has already been explained that MRA are subject to heteroskedasticity because *e-s* estimates are obtained with varying degrees of precision. That is to say, their respective standard errors are not the same. In economic sciences, e-s estimates correspond to partial regression coefficients drawn from primary studies. When estimating these coefficients, primary studies also estimate their standard errors. These estimates provide a measure of the MRA heteroskedasticity. This information may be used to adequately correct for heteroskedasticity. The Weighted Least-Squares (WLS) method of estimation takes such information explicitly into account in its estimation procedure.

The weighted least-squares method of estimation

A second alternative consists in estimating the parameters by using the WLS regression. Indeed, if y_i's variances are known, the most straightforward method of the correction of heteroskedasticity is by means of WLS[17].

Let $\sigma_{\varepsilon,i}$ be the estimated standard error[18] of the *i-th e-s* estimate, y_i, for any i. Knowing the y_i's heteroskedastic variances, $\sigma_{\varepsilon,i}^2$, the WLS method of estimation takes this information into account explicitly by, first, dividing Eq. (3) by the standard errors of y_i, $\sigma_{\varepsilon,i}$, giving:

$$\frac{y_i}{\sigma_{\varepsilon,i}} = \alpha \cdot \frac{1}{\sigma_{\varepsilon,i}} + \sum_{l}^{K-1} \beta_l \frac{x_{l,i}}{\sigma_{\varepsilon,i}} + \frac{\varepsilon_i}{\sigma_{\varepsilon,i}}, \quad \forall i = 1, \ldots, I \tag{4}$$

Second, the Ordinary Least-Squares (OLS) method of estimation is applied to the transformed variables, i.e. to Eq. (4).

[15] A wider confidence interval of a coefficient, say β_i, means that its variance, $\sigma_{\beta_i}^2$, is greater than expected. Thus, it conducts to a decrease of the *t-value* of β_i, $t_{\beta_i} = \hat{\beta}_i / \sqrt{\sigma_{\beta_i}^2}$, which increases the probability of falsely accepting the null hypothesis of tests of significance.

[16] See for instance Nelson and Kennedy [92] for a review of heteroskedasticity treatments used in meta-analysis studies dealing with environmental economics issues.

[17] As explained in Gujarati [116], once the original model has been transformed, the variance of "new" disturbance terms, ε_i^*, is:

$$\begin{aligned} Var(\varepsilon_i^*) &= E(\varepsilon_i^{*2}) = E\left(\left(\frac{\varepsilon}{\sigma_{\varepsilon,i}}\right)^2\right) \\ &= \frac{1}{\sigma_{\varepsilon,i}^2} E(\varepsilon) \quad \text{since} \sigma_{\varepsilon,i}^2 \text{ is known} \\ &= \frac{1}{\sigma_{\varepsilon,i}^2} (\sigma_{\varepsilon,i}^2) \quad \text{since} E(\varepsilon) = \sigma_{\varepsilon,i}^2 \\ &= 1 \end{aligned}$$

which is a constant. That is, the variance of the transformed error term, ε_i^*, is now homoskedastic.

[18] Again, like *e-s* estimates, estimated standard errors are drawn from primary studies.

The more $\sigma_{\varepsilon,i}$ is important, the less is the precision of y_i. Thus, by dividing each y_i by its standard error estimate, $\sigma_{\varepsilon,i}$, the WLS allocates to each *e-s* estimate a weight which is inversely proportional to its degree of precision. Intuitively, less precise *e-s* estimates, y_i with wider $\sigma_{\varepsilon,i}$, obtain relatively smaller weight than more precise ones in minimizing the (weighted) sum of residual squares. Indeed, recall that the OLS method consists of minimizing the sum of residual squares:

$$Min \sum_{i=1}^{I} e_i^2 = Min\left(\underset{(1,I)(I,1)}{e' \; e} \right)$$

where $e_{(I,1)}$ is the column vector of residuals defined as follows:

$$\begin{aligned} Y_{(I,1)} &= X_{(I,K)} \underset{(K,1)}{\hat{\beta}} + e_{(I,1)} \\ \Leftrightarrow \quad e_{(I,1)} &= Y_{(I,1)} - X_{(I,K)} \underset{(K,1)}{\hat{\beta}} \end{aligned}$$

where $\beta_{(K,1)}$ is the column vector of parameters estimated by the OLS method.

Thus, applying the OLS method to Eq. (4), WLS parameters estimates are obtained by minimizing:

$$Min \sum_{i=1}^{I} \left(\frac{e_i}{\sigma_{\varepsilon,i}} \right)^2$$

$$\Leftrightarrow \quad Min \sum_{i=1}^{I} \frac{1}{\sigma_{\varepsilon,i}^2} \cdot e_i^2 \tag{5}$$

$$\Leftrightarrow \quad Min \sum_{i=1}^{I} w_i e_i^2 \tag{6}$$

According to Eqs. (5) and (6), the WLS estimators are obtained by minimizing a weighted sum of residual squares with the y_is unconditional variances acting as the weights[19] :

$$w_i = \frac{1}{Var(y_i)} \tag{7}$$

Weights defined in Eq. (7) are known as being those that minimize the variance of the WLS estimators. These weights will then provide estimators that are BLUE (Best Linear Unbiased Estimators). In a particular framework of MA (the Fixed Effects Size model), these particular weights are obtained from the estimated standard error of each *e-s* estimates, y_i, drawn directly from primary studies [93,108].

Appendix C. Supplementary Information

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.rser.2013.04.021.

References

[1] International Energy Agency (IEA). World energy outlook 2012. OECD Publishing: 2012.
[2] Eide A. The right to food and the impact of liquid biofuels (Agrofuels): 2008.
[3] Benoist A, Dron D, Zoughaib A. Origins of the debate on the life-cycle greenhouse gas emissions and energy consumption of first-generation biofuels—a sensitivity analysis approach. Biomass and Bioenergy 2012;40: 133–42.

[19] The comparison of Eq. (4) and Eq. (6) may explain some confusion encountered in the literature. As highlighted by Nelson and Kennedy, some studies refer to weights based on variances and others refer to weights based on standard errors. As shown in Eq. (4), standard errors weights, $\sigma_{\varepsilon,i}$, are used to transform the variables, but as a consequence (see eq. (5)), it is variances weights, $\sigma_{\varepsilon,i}^2$, which are required to minimize the weighted sum of residual squares.

[4] Williams PRD, Inman D, Aden A, Heath GA. Environmental and sustainability factors associated with next-generation biofuels in the US: what do we really know? Environmental Science & Technology 2009;43:4763–75.
[5] European Parliament. Directive 2009/28/EC of the European Parliament and of the Council of 23 April 2009 on the promotion of the use of energy from renewable sources and amending and subsequently repealing Directives 2001/77/EC and 2003/30/EC. Official Journal of the European Union 2009; L 140/16: p. 16–62.
[6] United States Environmental Protection Agency (EPA), USEPA. Renewable Fuel Standard Program (RFS2) regulatory impact analysis: 2010.
[7] Global Biofuels Center. Special Report—next generation biofuels: policy review and outlook: 2010.
[8] Kaltschmitt M, Reinhardt GA, Stelzer T. Life cycle analysis of biofuels under different environmental aspects. Biomass and Bioenergy 1997;12:121–34.
[9] Menichetti E, Otto M. Energy balance and greenhouse gas emissions of biofuels from a life-cycle perspective. In: Howarth RW, Bringezu S, editors. Biofuels: environmental consequences and interactions with changing land use. Proceedings of the Scientific Committee on Problems of the Environment (SCOPE) international biofuels project rapid assessment, Ithaca NY, University of Cornell; 2008: p. 81—109.
[10] Gnansounou E, Dauriat A, Villegas J, Panichelli L. Life cycle assessment of biofuels: energy and greenhouse gas balances. Bioresource Technology 2009;100:4919–30.
[11] Cherubini F, Bird ND, Cowie A, Jungmeier G, Schlamadinger B, Woess-Gallasch S. Energy- and greenhouse gas-based LCA of biofuel and bioenergy systems: key issues, ranges and recommendations. Resources, Conservation and Recycling 2009;53:434–47.
[12] Sunde K, Brekke A, Solberg B. Environmental impacts and costs of hydrotreated vegetable oils, transesterified lipids andwoody BTL—a review. Energies 2011;4:845–77.
[13] Borrion AL, McManus MC, Hammond GP. Environmental life cycle assessment of lignocellulosic conversion to ethanol: a review. Renewable and Sustainable Energy Reviews 2012;16:4638–50.
[14] Wiloso EI, Heijungs R, De Snoo GR. LCA of second generation bioethanol: a review and some issues to be resolved for good LCA practice. Renewable and Sustainable Energy Reviews 2012;16:5295–308.
[15] Brandão M, Heath G, Cooper J. What can meta-analyses tell us about the reliability of life cycle assessment for decision support? Journal of Industrial Ecology 2012;16:S3–7.
[16] Heath GA, Mann MK. Background and reflections on the life cycle assessment harmonization project. Journal of Industrial Ecology 2012;16:S8–11.
[17] Handler RM, Canter CE, Kalnes TN, Lupton FS, Kholiqov O, Shonnard DR, et al. Evaluation of environmental impacts from microalgae cultivation in open-air raceway ponds: analysis of the prior literature and investigation of wide variance in predicted impacts. Algal Research 2012;1:83–92.
[18] Liu X, Clarens AF, Colosi LM. Algae biodiesel has potential despite inconclusive results to date. Bioresource Technology 2012;104:803–6.
[19] Abreu M, De Groot HLFM, Florax RJGM. A meta-analysis of β-convergence: the legendary two-percent. The Journal of Economic Surveys 2005;19: 389–420.
[20] International Organization for Standardization (ISO). ISO14040:2006. Environmental management-life cycle assessment—principles and framework: 2006.
[21] International Organization for Standardization (ISO). ISO14044:2006 Environmental management—life cycle assessment—requirements and guidelines: 2006.
[22] European Commission—Joint Research Centre—Institute for Environment and Sustainability. International Reference Life Cycle Data System (ILCD) Handbook—general guide for life cycle assessment—detailed guidance. Luxembourg; Publications Office of the European Union: 2010.
[23] Earles JM, Halog A. Consequential life cycle assessment: a review. The International Journal of Life Cycle Assessment 2011;16:445–53.
[24] Weidema BP, Frees N, Nielsen A-M. Marginal production technologies for life cycle inventories. The International Journal of Life Cycle Assessment 1999;4:48–56.
[25] Tillman A-M, Ekvall T, Baumann H, Rydberg T. Choice of system boundaries in life cycle assessment. Journal of Cleaner Production 1994;2:21–9.
[26] Vigon BW et al. Life-cycle assessment: inventory guidelines and principles. EPA/600/R-ed. Washington, DC, 20460: 1993.
[27] Searchinger T, et al. Use of US croplands for biofuels increases greenhouse gases through emissions from land-use change. Science 2008;319: 1238–40.
[28] Crutzen PJ, Mosier AR, Smith KA, Winiwarter W. N_2O release from agro-biofuel production negates global warming reduction by replacing fossil fuels. Atmospheric Chemistry and Physics Discussions 2007;7:11191–205.
[29] Bio Intelligence Service. Analyses de cycle de vie appliquées aux biocarburants de première génération consommés en France—Rapport Final; Paris: 2010.
[30] Houghton J et al., editor. Revised 1996 IPCC Guidelines for national greenhouse gas inventories: workbook—module 24 agriculture: 1996.
[31] Farrell AE, Plevin RJ, Turner BT, Jones AD, O'Hare M, Kammen DM. Ethanol can contribute to energy and environmental goals. Science 2006;311: 506–8.
[32] Whittaker C, McManus MC, Hammond GP. Greenhouse gas reporting for biofuels: a comparison between the RED, RTFO and PAS2050 methodologies. Energy Policy 2011;39:5950–60.

[33] Warner ES, Heath GA. Life cycle greenhouse gas emissions of nuclear electricity generation. Journal of Industrial Ecology 2012;16:S73–92.
[34] Burkhardt JJ, Heath G, Cohen E. Life cycle greenhouse gas emissions of trough and tower concentrating solar power electricity generation. Journal of Industrial Ecology 2012;16:S93–109.
[35] Kim HC, Fthenakis V, Choi J-K, Turney DE. Life cycle greenhouse gas emissions of thin-film photovoltaic electricity generation. Journal of Industrial Ecology 2012;16:S110–21.
[36] Hsu DD, et al. Life cycle greenhouse gas emissions of crystalline silicon photovoltaic electricity generation. Journal of Industrial Ecology 2012;16:S122–35.
[37] Dolan SL, Heath GA. Life cycle greenhouse gas emissions of utility-scale wind power. Journal of Industrial Ecology 2012;16:S136–54.
[38] Bureau J-C, Disdier A-C, Gauroy C, Tréguer D. A quantitative assessment of the determinants of the net energy value of biofuels. Energy Policy 2010;38:2282–90.
[39] Glass GV. Primary, secondary, and meta-analysis of research. Educational Researcher 1976;5:3–8.
[40] Glass GV. Integrating findings: the meta-analysis of research. Review of Research in Education 1977;5:351–79.
[41] Glass GV, McGaw B, Smith ML. Meta-analysis in social research. Thousand Oaks, CA: Sage Publications; 1981.
[42] Stanley TD, Jarrell SB. Meta-regression analysis: a quantitative method of literature surveys. The Journal of Economic Surveys 1989;3:161–70.
[43] Capelle-Blancard G, Couderc N. How do shareholders respond to downsizing? A meta-analysis: 2007.
[44] Stanley TD. Wheat from chaff: meta-analysis as quantitative literature review. Journal of Economic Perspectives 2001;15:131–50.
[45] Bai Y, Luo L, Voet E. Life cycle assessment of switchgrass-derived ethanol as transport fuel. The International Journal of Life Cycle Assessment 2010;15: 468–77.
[46] Batan L, Quinn J, Willson B, Bradley T. Net energy and greenhouse gas emission evaluation of biodiesel derived from microalgae. Environmental Science & Technology 2010;44:7975–80.
[47] Campbell PK, Beer T, Batten D. Life cycle assessment of biodiesel production from microalgae in ponds. Bioresource Technology 2010;102:50–6.
[48] Cherubini F, Strømman AH, Ulgiati S. Influence of allocation methods on the environmental performance of biorefinery products—a case study. Resources, Conservation and Recycling 2011;55:1070–7.
[49] Choudhury R, Wurster R, Weber T. GM Well-to-Wheel analysis of energy use and greenhouse gas emissions of advanced fuel/vehicle systems—a European study: 2002.
[50] Chouinard-Dussault P, Bradt L, Ponce-Ortega JM, El-Halwagi MM. Incorporation of process integration into life cycle analysis for the production of biofuels. Clean Technologies and Environmental Policy 2010;13:673–85.
[51] Delucchi MA. Lifecycle Analyses of Biofuels: 2006.
[52] Elsayed MA, Matthews R, Mortimer ND. Carbon and Energy Balances for a Range of Biofuels Options: 2003.
[53] Fazio S, Monti A. Life cycle assessment of different bioenergy production systems including perennial and annual crops. Biomass and Bioenergy 2011;35:4868–78.
[54] González-García S, Gasol CM, Gabarrell X, Rieradevall J, Moreira MT, Feijoo G. Environmental profile of ethanol from poplar biomass as transport fuel in Southern Europe. Renewable Energy 2010;35:1014–23.
[55] González-García S, Moreira MT, Feijoo G. Environmental performance of lignocellulosic bioethanol production from Alfalfa stems. Biofuels, Bioproducts and Biorefining 2010;4:118–31.
[56] Groode TA, Heywood JB. Ethanol: A Look Ahead: 2007.
[57] Haase M, Skott S, Fröhling M. Ecological evaluation of selected 1st and 2nd generation biofuels—FT fuel from wood and ethanol from sugar beets. In: Hiete M, Ludwig J, Bidart C, Schultmann FH, editors. Challenges for sustainable biomass utilisation. Proceedings of the Chilean–German biociclo workshop, Karlsruhe, 26.03.2009; KIT Scientific Publishing: 2010.
[58] Hoefnagels R, Smeets E, Faaij A. Greenhouse gas footprints of different biofuel production systems. Renewable and Sustainable Energy Reviews 2010;14:1661–94.
[59] Hsu DD, Inman D, Heath GA, Wolfrum EJ, Mann MK, Aden A. Life cycle environmental impacts of selected US ethanol production and use pathways in 2022. Environmental Science & Technology 2010;44:5289–97.
[60] JEC—Joint Research Centre-EUCAR-CONCAWE collaboration. Well-to-wheels analysis of future automotive fuels and powertrains in the European context (Version 3c): 2011.
[61] JEC—Joint Research Centre-EUCAR-CONCAWE collaboration. Well-to-wheels analysis of future automotive fuels and powertrains in the European context (Version 2c): 2007.
[62] Jungbluth N, Frischknecht R, Emmenegger MF, Steiner R, Tuschmid M. Life cycle assessment of BTL-fuel production: life cycle impact assessment and interpretation. Renewable Fuels for Advanced Powertrains (RENEW) Integrated Project. Uster, Switzerland, ESU-services: 2007.
[63] Jungbluth N, Büsser S, Frischknecht R, Tuchschmid M. Life cycle assessment of biomass-to-liquid fuels. Uster, Switzerland: 2008.
[64] Kaufman AS, Meier PJ, Sinistore JC, Reinemann DJ. Applying life-cycle assessment to low carbon fuel standards—how allocation choices influence carbon intensity for renewable transportation fuels. Energy Policy 2010;38:5229–41.
[65] Koponen K, Soimakallio S, Sipilä E. Assessing the greenhouse gas emissions of waste-derived ethanol in accordance with the EU RED methodology for biofuels.Vuorimiehentie. Finland: VTT Technical Research Center; 2009.
[66] Lardon L, Hélias A, Sialve B, Steyer J-P, Bernard O. Life-cycle assessment of biodiesel production from microalgae. Environmental Science & Technology 2009;43:6475–81.
[67] Luo L, Voet E, Huppes G. Udo de Haes HA. Allocation issues in LCA methodology: a case study of corn stover-based fuel ethanol. The International Journal of Life Cycle Assessment 2009;14:529–39.
[68] McKechnie J, et al. Impacts of co-location, co-production, and process energy source on life cycle energy use and greenhouse gas emissions of lignocellulosic ethanol. Biofuels, Bioproducts and Biorefining 2011;5:279–92.
[69] Mehlin M, Reichmuth M, Wacker M. Renewable fuels for cross border transportation: 2003.
[70] Mu D, Seager T, Rao PS, Zhao F. Comparative life cycle assessment of lignocellulosic ethanol production: biochemical versus thermochemical conversion. Environmental Management 2010;46:565–78.
[71] Mullins KA, Griffin WM, Matthews HS, States U. Policy implications of uncertainty in modeled life-cycle greenhouse gas emissions of biofuels. Environmental Science & Technology 2011;45:132–8.
[72] Sander K, Murthy GS. Life cycle analysis of algae biodiesel. The International Journal of Life Cycle Assessment 2010;15:704–14.
[73] Schmitt E, Bura R, Gustafson R, Cooper J, Vajzovic A. Converting lignocellulosic solid waste into ethanol for the State of Washington: an investigation of treatment technologies and environmental impacts. Bioresource Technology 2012;104:400–9.
[74] Sheehan J, Aden A, Paustian K, Brenner J, Walsh M, Nelson R. Energy and environmental aspects of using corn stover for fuel ethanol. Journal of Industrial Ecology 2004;7:117–46.
[75] Spatari S, Bagley DM, MacLean HL. Life cycle evaluation of emerging lignocellulosic ethanol conversion technologies. Bioresource Technology 2010;101:654–67.
[76] Spatari S, MacLean HL. Characterizing model uncertainties in the life cycle of lignocellulose-based ethanol fuels. Environmental Science & Technology 2010;44:8773–80.
[77] Spatari S, Zhang Y, MacLean HL. Life cycle assessment of switchgrass-and corn stover-derived ethanol-fueled automobiles. Environmental Science & Technology 2005;39:9750–8.
[78] Stephenson AL, Dupree P, Scott SA, Dennis JS. The environmental and economic sustainability of potential bioethanol from willow in the UK. Bioresource Technology 2010;101:9612–23.
[79] Stephenson AL, Kazamia E, Dennis JS, Howe CJ, Scott SA, Smith AG. Life-cycle assessment of potential algal biodiesel production in the United Kingdom: a comparison of raceways and air-lift tubular bioreactors. Energy & Fuels 2010;24:4062–77.
[80] Stichnothe H, Azapagic A. Bioethanol from waste: life cycle estimation of the greenhouse gas saving potential. Resources, Conservation and Recycling 2009;53:624–30.
[81] Stratton RW, Wong HM, Hileman JI. Life cycle greenhouse gas emissions from alternative jet fuels, PARTNER Project 28 Report Version 1.1; Cambridge, Massachusetts: 2010.
[82] Van Vliet OPR, Faaij APC, Turkenburg WC. Fischer–Tropsch diesel production in a well-to-wheel perspective: a carbon, energy flow and cost analysis. Energy Conversion and Management 2009;50:855–76.
[83] Vera-morales M, Schäfer A. Final Report: fuel-cycle assessment of alternative aviation fuels: 2009.
[84] Wang M, Han J, Haq Z, Tyner WE, Wu M, Elgowainy A. Energy and greenhouse gas emission effects of corn and cellulosic ethanol with technology improvements and land use changes. Biomass and Bioenergy 2011;35: 1885–96.
[85] Wang M, Huo H, Arora S. Methods of dealing with co-products of biofuels in life-cycle analysis and consequent results within the US context. Energy Policy 2011;39:5726–36.
[86] Wu M, Wu Y, Wang M Mobility chains analysis of technologies for passenger cars and light-duty vehicles fueled with biofuels: application of the GREET model to the Role of Biomass in America's Energy Future (RBAEF) project: 2005.
[87] Xie X, Wang M, Han J. Assessment of fuel-cycle energy use and greenhouse gas emissions for Fischer–Tropsch diesel from coal and cellulosic biomass. Environmental Science & Technology 2011;45:3047–53.
[88] Parry M, Canziani O, Palutikof J, Van der Linden P, Hanson C, editors. Climate change 2007: impacts, adaptation and vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge, UK; Cambridge University Press: 2007.
[89] Whitaker M, Heath GA, O'Donoughue P, Vorum M. Life cycle greenhouse gas emissions of coal-fired electricity generation. Journal of Industrial Ecology 2012;16:S53–72.
[90] JEC—Joint Research Centre-EUCAR-CONCAWE collaboration. Well-to-Wheels analysis of future automotive fuels and powertrains in the European context (Version 3). JRC/EUCAR/CONCAWE European Commission Joint Research Center: 2008.
[91] Melamu R, Von Blottnitz H. 2nd generation biofuels a sure bet? A life cycle assessment of how things could go wrong Journal of Cleaner Production 2011;19:138–44.
[92] Nelson JP, Kennedy PE. The use (and abuse) of meta-analysis in environmental and natural resource economics: an assessment: 2008.
[93] Hedges LV, Olkin I. Statistical methods for meta-analysis. New York: Academic Press; 1985.

[94] Florax RJGM, Poot J. Learning from the flood of numbers: meta-analysis in economics: 2005.
[95] White H. A heteroscedasticity-consistent covariance matrix estimator and a direct test for heteroscedasticity. Econometrica 1980;48:817–38.
[96] McLaughlin SB, Walsh ME. Evaluating environmental consequences of producing herbaceous crops for bioenergy. Biomass and Bioenergy 1998;14: 317–24.
[97] Aden A et al. Lignocellulosic biomass to ethanol process design and economics utilizing co-current dilute acid prehydrolysis and enzymatic hydrolysis for corn stover. Golden, Colorado: 2002.
[98] Aatola H, Larmi M, Sarjovaara T. Hydrotreated Vegetable Oil (HVO) as a renewable diesel fuel: trade-off between NO*x*, particulate emission, and fuel consumption of a heavy duty engine. Helsinki: 2008.
[99] Jorquera O, Kiperstok A, Sales EA, Embiruçu M, Ghirardi ML. Comparative energy life-cycle analyses of microalgal biomass production in open ponds and photobioreactors. Bioresource Technology 2010;101:1406–13.
[100] Brennan L, Owende P. Biofuels from microalgae—a review of technologies for production, processing, and extractions of biofuels and co-products. Renewable and Sustainable Energy Reviews 2010;14:557–77.
[101] Mata TM, Martins Aa, Caetano NS. Microalgae for biodiesel production and other applications: a review. Renewable and Sustainable Energy Reviews 2010;14:217–32.
[102] Brander M, Tipper R, Hutchison C, Davis G. Consequential and attributional approaches to LCA: a guide to policy makers with specific reference to greenhouse gas LCA of biofuels. Ecometrica Press 2009;44:1–14.
[103] IEA. Bioenergy ExCo63 Workshop. Bioenergy—the impact of indirect land use change: 2009.
[104] Dwivedi P, Alavalapati JRR, Lal P. Cellulosic ethanol production in the United States: conversion technologies, current production status, economics, and emerging developments. Energy for Sustainable Development 2009;13:174–82.
[105] Cheng J. Biomass to renewable energy processes. London, UK: CRC Press Inc.; 2010.
[106] Hamelinck CN, Faaij APC. Outlook for advanced biofuels. Energy Policy 2006;34:3268–83.
[107] Pienkos P, Darzins A. The promise and challenges of microalgal-derived biofuels. Biofuels, Bioproducts and Biorefining 3, 2009:431–40.
[108] Hedges LV. Fixed effects models. In: Cooper H, Hedges LV, editors. The handbook of research synthesis. New York: Russell Sage Foundation; 1994 p.. 285–99.
[109] Jarrell SB, Stanley TD. A meta-analysis of the union–nonunion wage gap. Industrial & Labor Relations Review 1990;44:54–67.
[110] Smith VK, Kaoru Y. Signals or noise? Explaining the variation in recreation benefit estimates American Journal of Agricultural Economics 1990;72: 419–33.
[111] Smith VK, Kaoru Y. What have we learned since Hotelling's letter? A meta-analysis Economics Letters 1990;32:277–81.
[112] Walsh RG, Johnson DM, McKean JR. Issues in nonmarket valuation and policy applications: a retrospective glance. Western Journal of Agricultural Economics 1989;14:178–88.
[113] Weitzman ML, Kruse DL. Profit sharing and productivity. In: Blinder A, editor. Paying for productivity: a look at the evidence. Washington, DC: Brookings Institution; 1990 p. 95–140.
[114] Hedges LV. Meta-analysis. Journal of Educational Statistics 1992;17:279–96.
[115] Lipsey MW, Wilson DB. Practical meta-analysis. Thousands Oaks, CA: Sage Publications; 2001.
[116] Gujarati DN. Basic econometrics. Fourth. New York, NY: The McGraw-Hill Companies Inc; 2004.
[117] Ready RC. Do landfills always depress nearby property values? 2005.
[118] Thompson SG, Sharp SJ. Explaining heterogeneity in meta-analysis: a comparison of methods. Statistics in Medicine 1999;18:2693–708.
[119] Dones R et al. Life cycle inventories of energy systems: results for current systems in Switzerland and other UCTE Countries Data v2.0 ecoinvent report No. 5: 2007.
[120] S. González-García, L. Luo, M.T. Moreira, G. Feijoo and G. Huppes, Life cycle assessment of flax shives derived second generation ethanol fueled automobiles in Spain, Renewable and Sustainable Energy Reviews 13, 2009, 1922–1933. http://dx.doi.org/10.1016/j.rser.2009.02.003.